

|             |   |
|-------------|---|
| Title       | 潜在的評価構造の差異を考慮した離散型選択モデル(Dissertation_全文)  |
| Author(s)   | 佐々木, 邦明   |
| Citation    | Kyoto University (京都大学)   |
| Issue Date  | 1998-03-23  |
| URL         | <a href="http://dx.doi.org/10.11501/3135644">http://dx.doi.org/10.11501/3135644</a> |
| Right       |   |
| Type        | Thesis or Dissertation  |
| Textversion | author  |

②

# 潜在的評価構造の差異を考慮した 離散型選択モデル

1997年12月

佐々木 邦明

# 潜在的評価構造の差異を考慮した離散型選択モデル

## 概要

非集計離散型選択モデルの導入当初は、その背景にある行動理論の明確さ故、時間的・地域的移転性や、交通発生や経路選択などの様々なレベルの問題に適用が可能であると考えられていた。しかし、これまでの様々な事例研究により、当初期待されていたほどの移転性や、様々なレベルでの予測に適用するのは難しいことが指摘されてきた。この原因と考えられるものの中で、特に主要な要因として、モデルに用いられなかった変数の問題、個人間の様々な異質性の問題、行動の時間的依存性といった動的特性の問題があげられる。そこで本研究は、これらの非集計行動モデルが抱えている問題点に焦点を当て、選択行動理論を含めて再構築を行い、それに即した非集計行動モデルの開発を行った。具体的には以下にあげる3つの点に着目している。

- ・主観的意識要因を考慮した非集計離散型選択モデルの構築
- ・個人の異質性を考慮した非集計離散型選択モデルの構築
- ・行動の動的特性を考慮した選択モデルの構築

これらの要因を考慮するため、マーケティングの分野で用いられている意思決定過程を背景となる選択行動理論として援用し、上記の3点について、多変量解析手法や生存時間モデルなどの、主に他の分野で使われている手法と非集計行動モデルの複合モデルを提案し、事例研究を通じてその有効性を確認する。

まず、主観的意識要因を明示的にモデル中に取り込むために、線形構造方程式モデルを用いて、主観的意識を尋ねたデータと客観的属性との間の関係を定式化した。そして、行動の背景となる潜在変数を導出し、その変数を非集計離散型選択モデルの説明変数に用いる手法を提案する。事例研究の結果から、提案した手法を用いることにより選択モデルの適合度を改善され、その有効性が示された。

同様に消費者意思決定過程に基づいて、線形構造方程式モデルを用いて個人の主観的な価値基準を尋ねたデータと客観的属性との関係を構造化し、主観的価値基準を表す潜在変数を導出した。この潜在変数を、離散型選択モデルの未知母数をノンパラメトリックに分布させる潜在セグメントモデルに用いて、個人の異質性を考慮した選択モデルを提案した。この手法の事例研究を仮想の状況下での選択を尋ねるS P (Stated Preference) データ、実際の選択結果を尋ねたR P

「Revealed Preference」データ及び同一個人に繰り返し調査を行うパネルデータに適用した結果、いずれも高い適合度を示し、この手法の有効性も確認された。

行動の動的特性を考慮するために、パネルデータを用いた非集計行動モデルの構築を行った。このとき、パネルデータ特有の問題であるパネルサンプルの消耗によるバイアスを修正するために、生存時間モデルを用いてサンプルに重みを付けるモデルを提案した。この手法の事例研究を行う際に、個人の時系列的な誤差の相関を考慮するための系列相関モデルを同時に用いたモデルも推定し、サンプル消耗によるパラメータバイアスを示しただけでなく、系列相関とサンプル消耗の関係を実証的に明らかにした。

## 目次

|       |                                   |    |
|-------|-----------------------------------|----|
| 第 1 章 | 序論                                | 1  |
| 1-1   | 本研究の背景と目的                         | 1  |
| 1-2   | 交通行動分析に関する研究の流れと今後の方向性            | 4  |
| 1-3   | 本論文の構成                            | 11 |
| 第 1 章 | 参考文献                              | 13 |
| 第 2 章 | 主観的意識要因を考慮した交通行動分析                | 15 |
| 2-1   | はじめに                              | 15 |
| 2-2   | 主観的潜在要因と数理心理学データの特性               | 16 |
| 2-2-1 | 潜在意識要因を交通行動分析に取り入れた既存研究のレビュー      | 16 |
| 2-2-2 | 交通機関選択に影響する潜在要因                   | 19 |
| 2-2-3 | 認知的不協和とカテゴリカルデータの特性               | 20 |
| 2-3   | モデルの定式化                           | 23 |
| 2-3-1 | 潜在変数モデル                           | 23 |
| 2-3-2 | LISREL モデル                        | 23 |
| 2-3-3 | フレームワーク                           | 26 |
| 2-4   | 推定方法の検討と数理心理データの特性を考慮した定式化        | 29 |
| 2-4-1 | 段階推定                              | 29 |
| 2-4-2 | 認知的不協和を考慮した段階推定                   | 31 |
| 2-4-3 | 同時推定                              | 33 |
| 2-4-4 | 知覚値指標の離散性を考慮した同時推定                | 34 |
| 2-5   | 事例研究                              | 36 |
| 2-5-1 | データの概略                            | 36 |
| 2-5-2 | モデルの特定化                           | 37 |
| 2-5-3 | 各推定結果の考察                          | 40 |
| 2-5-4 | 推定結果の比較                           | 45 |
| 2-6   | 本章のまとめ                            | 49 |
| 第 2 章 | 参考文献                              | 52 |
| 第 3 章 | 個人の異質性を考慮した交通行動分析                 | 55 |
| 3-1   | はじめに                              | 55 |
| 3-2   | 個人の異質性を考慮した既存研究のレビュー              | 57 |
| 3-3   | 個人に繰り返し尋ねた S P データを用いた個人ベースの行動モデル | 61 |
| 3-3-1 | 個人から繰り返し採取した S P データの性質について       | 61 |
| 3-3-2 | 複数 S P データによる個人パラメータの推定           | 63 |
| 3-3-3 | 個人パラメータに基づいたセグメンテーション             | 65 |



|         |                             |     |
|---------|-----------------------------|-----|
| 3-4     | 個人の潜在的態度を考慮した行動モデル          | 70  |
| 3-4-1   | 個人の態度と選択特性                  | 70  |
| 3-4-2   | 態度データの特性とその定性的分析            | 71  |
| 3-4-3   | L I S R E Lモデルによる態度データの分析   | 73  |
| 3-4-4   | 態度変数によるセグメンテーション            | 75  |
| 3-4-5   | 潜在クラス分析と潜在セグメント分析           | 81  |
| 3-4-6   | 態度変数を用いた潜在セグメント分析           | 86  |
| 3-5     | 本章のまとめ                      | 92  |
| 第3章     | 参考文献                        | 95  |
| 4章      | 個人の異質性を考慮したパネル分析            | 99  |
| 4-1-1   | 本章の背景と目的                    | 99  |
| 4-1-2   | パネルデータの特徴とパネルデータを用いた研究のレビュー | 101 |
| 4-1-2-1 | パネルデータの特徴                   | 102 |
| 4-1-2-2 | パネルデータを用いた既存研究のレビュー         | 102 |
| 4-1-3   | パネルデータを用いた潜在セグメント分析         | 107 |
| 4-1-3-1 | モデルのフレームワーク                 | 107 |
| 4-1-3-2 | 用いるデータの概要                   | 109 |
| 4-1-3-3 | 客観的属性を用いた潜在セグメント分析          | 111 |
| 4-1-3-4 | 潜在的態度を用いた潜在セグメント分析          | 115 |
| 4-1-4   | 本章のまとめ                      | 124 |
| 第4章     | 参考文献                        | 126 |
| 5章      | パネル消耗を考慮した選択モデル分析           | 129 |
| 5-1     | 本章の背景と目的                    | 129 |
| 5-1-2   | パネル消耗の特徴と既存研究のレビュー          | 130 |
| 5-1-2-1 | パネル消耗の特質                    | 130 |
| 5-1-2-2 | パネル消耗に関する既存研究のレビュー          | 132 |
| 5-1-2-3 | 調査費協力層の存在とサンプルバイアス          | 136 |
| 5-1-3   | 消耗効果を考慮した選択モデル分析            | 138 |
| 5-1-3-1 | パネル調査における消耗の概要              | 138 |
| 5-1-3-2 | 消耗バイアスの定性的・定量的分析            | 140 |
| 5-1-3-3 | 確率的消耗モデルを用いたパネル分析           | 146 |
| 5-1-3-4 | パネル消耗によるバイアスを修正した買物目的地選択モデル | 153 |
| 5-1-4   | 本章のまとめ                      | 156 |
| 第5章     | 参考文献                        | 158 |
| 第6章     | 研究のまとめと今後の課題                | 161 |

## 第1章 序論

### 1-1 本研究の背景と目的

大都市での慢性化した交通渋滞、その一方での朝夕の通勤電車の混雑、観光地周辺道路での休日の大渋滞とそれに伴う交通混雑や生活道路への通過交通の進入などの地域住民の生活環境の悪化、地方都市の郊外部での慢性的渋滞と公共交通の衰退、交通に起因する環境汚染やエネルギー消費の増加、このようにいくつもの解決すべき問題を抱えているのが、今日のわが国の交通の現状であるといえる。これらの問題はその社会的非効率性という共通の背景を抱えている。社会的非効率性とは経済学で言う「パレート最適」が達成されていない、つまり誰の効用も下げずに全体の効用をあげる資源配分が、可能な状態であるということである。これは渋滞や環境汚染などの外部不経済の発生という、技術的・制度的な問題による市場の失敗に起因するものであり、適切な社会基盤整備や交通政策によってパレート最適が達成されうると考えられる。そのための具体的な政策として、パーク＆ライドなどを用いた交通需要マネジメント政策が、この数年わが国でも深刻な交通渋滞に悩むいくつかの地方都市で試験的に実施されている。また大都市圏では、自動化運転、経路案内などを統合したI T S（Intelligent Transport Systems）が、ビーコンなどを用いて情報提供を行うV I C S（Vehicle Information and Communication Systems）の供用が始まるなど実用化の段階に入り、これらを用いて経路需要を含めた総合的な交通需要の管理方策が検討されている。これまでに試験的に運用された交通需要マネジメントの事後分析を見ると、導入された交通需要マネジメント政策に対する各個人の対応は、すぐに行動に変化が現れる人や、徐々に行動を変化させる人、ほとんど変化しない人といった反応時間の違いを含めて、様々な行動の変化のパターンが観察されている。これらを詳しく見ると、その意思決定には個人の特性、時刻、状況など非常に多くの要素が絡み合っていることがうかがえる。例えば通勤時の交通機関選択に対する交通需要マネジメント政策としてのパーク＆バスライドの利用状況を例にとると、その日の残業の有無などの勤務状況に依存し、終業後の予定にも応じて変動するなど、決して料金と所要時間のトレードオフといったシンプルな意思決定メカニズムに従っているとは言い難い。同じように交通需要マネジメント政策の一つであるロードプライシングを、外部費用の内部化による市場価格の修正として分析する場合に、需要関数を導出するための交通行動モデルは、こういった複雑な行動特性を考慮で

きるものであることが望ましい。

これまで、交通需要モデルにはマクロ的・全体的観点から分析を行う、いわゆる4段階推定法に代表される集計分析、ミクロ的・個人的観点から分析を行う非集計分析の2系統が存在している。誕生してから約20年足らずの非集計分析も、現在では実務レベルへの浸透も進み、これら2系統の手法を適切な場面に応じての使い分けがなされはじめている。ここでモデルとは、本来的にどのような特性を備えているべきであるかを考えてみる。モデルは、ある現象を単純化して描写したものであり、現実世界の特性をいくつか備えてはいるが、総てを備えている必要はないこともある。例えば交通現象を取り扱う場合でも、総てを説明できるモデルを構築することは現時点であり現実的ではなく、そのときの分析対象によっては、シンプルな形のモデルによる場合が好ましいことも多い。しかし、先に述べたような交通需要マネジメント政策や交通情報提供などの効果分析を行う場合の交通行動モデルは、個人による反応の違いが顕著であり、その反応の違いが政策の効果に大きな影響を与えるので、個人の選択行動特性の違いをより細かく表すことのできる非集計モデルが適していると考えられる。しかし、その非集計分析を行う場合、開発当時から用いられているような限られた、所要時間や費用などのいくつかの主要な要因に基づいて、効用最大化原理に基づいて行動する合理的個人の仮定や、対象母集団内では同一の嗜好を持つという仮定の下で静学的にモデルを構築したのでは、政策の効果을正しく計量することは、先に挙げた理由により困難であると思われる。効果を正確に計量できないモデルによる予測に基づいて政策が決定された場合は、期待された効果を発揮できないだけでなく、逆にパレート最適から遠ざかる可能性すら存在する。このように、より個人の行動を精緻に記述できるモデルは、経路所要時間をはじめとする交通情報が提供されることによって、本当に渋滞の緩和が行えるのか、という疑問に明確な答えを導くためには必要となるであろう。

自動車交通を利用する人が、現在の交通の問題点として真っ先に挙げるのは、大都市圏に限らず日々発生している交通渋滞であろう。これらの渋滞の多くは、交通量がわずかに減少するだけで大きく緩和されるという報告がいくつか見られる。そのため、フレックスタイムはもちろんパーク＆ライドなどの、これまでに提案され試験的に運用されている交通需要マネジメント政策を、政策の効果が大きく持続性のあるサブグループに重点的に働きかけることが、母集

団全体に働きかけて行くより効率的かつ効果的だという指摘がある。このとき、個人の異質性を表現できるモデルが構築され、そのような政策が効果的であると思われるグループを構成する人々の特徴が把握できるとしたら、その意義は非常に大きいと思われる。これらの指摘から、従来型の合理的個人が均質の嗜好を持つと仮定するアプローチでは、交通需要マネジメント政策の効果測定に、自ずと限界が見えてくるとと思われる。

また、同様なことから、これまでの単一断面でのデータに基づいた静学的モデルの限界もうかがえる。なぜならば、個人の交通行動の選択は、ある時点での決定やその結果が、後の時空間での交通行動の意思決定に影響を与えていることが明白な場合も多いためである。例えば、自家用車の購入と使用について考えると、世帯の自家用車によるトリップ数およびトリップ延長は、自家用車を購入したという、ある一時点での意思決定に依存することは明白である。もちろん自家用車の購入という意思決定に、その後の自家用車によるトリップによって得られる期待効用も影響を与える。このように自家用車の購入と自家用車によるトリップは、明らかに時点間で独立とはいいがたい。しかし、単一断面での調査に基づくモデルは、用いたサンプル間の差異によって行動の差異を説明するものであり、サンプルの変化によって行動の変化を説明するものではない。このことは、行動結果と調査時点の関係を独立と暗黙裏に仮定していることになる。例えば、一時点の断面調査データから構築された、自家用車の購入決定と自家用車によるトリップ数を説明するモデルは、説明される二つの変数の時間的な独立を仮定していることになる。この例のように、交通行動モデルは、サンプル間の差異だけではなく、サンプルの属性や交通条件などの時間的変化を、明示的に考慮できるものでなければならない場合も多く存在する。

そこで本研究は、先に挙げた個人の異質性や行動の時間的依存性などの、交通需要予測のベースとなる交通行動モデルが今後直面し解決していかなければならないと考えられる分析上の問題に対して、その回答となる手法の提案を試みるものである。そのためには、潜在的な個人の意思決定構造を、より積極的・明示的にモデリングすることが必要である。具体的には、交通機関選択に個人の潜在的な要因を取り込む手法、および、個人の主観的な嗜好の違いを明示した交通機関選択モデルの提案である。また、そのアプリケーションとして、嗜好の違いが現れやすい選好意識データ（SPデータ）や、個人の選択の時系列データであるパネルデータを

用いて、嗜好の違いを明示的に示すモデル推定する。また、パネルデータには特有の消耗現象によるバイアスが存在することが多いため、パネル調査への残留を内生的にモデル化した交通行動モデルの提案を行う。このように様々な要素が複雑に絡み合った交通行動を、計測が困難と思われる要素の意思決定への作用、個人ごとに生じている様々な嗜好の違い、そして時間的な要因を含めたモデルを構築し、交通行動のほんの一部の再現が可能になることを願うものである。

## 1-2 交通行動分析に関する研究の流れと今後の方向性

日本の交通需要予測の分析に、非集計離散型選択モデルが本格的に導入されるきっかけとなった「交通システム分析と交通計画」と題したセミナーが M. Manheim, M. Ben-Akiva, S. Lerman 氏を招いて 1981 年に行われた。その後、交通需要分析を行っている研究者の間で、従来からのパーソントリップ調査など、大規模調査に基づいた 4 段階推定法に代表される集計型の分析に置き換わるものとして期待され、国内でも多くの研究者により、その事例研究が盛んになった。非集計交通行動モデルは、1970 年代初頭にアメリカ合衆国で D. McFadden らによってその理論的体系が形作られ、欧米では日本に先駆けてそのアプリケーションが盛んに研究されていた。非集計行動モデルは従来からの集計型分析に比べ、

- ・小サンプルでモデルの推定が可能である。
- ・種々の政策変数の導入が容易である。
- ・選択行動を表すという汎用性の広い行動論的理論背景が明確である。

などの長所を持っていたため、交通需要予測のみならず、立地分析などにもその適用範囲を広げながら、わが国でも土木計画の分野全般に浸透していった。その一方、非集計モデルの開発の過程で指摘された、選択肢集合の不確実性の問題や意思決定者の異質性の問題は、国内では非集計モデル導入から方法論的な確立を見るまでの数年間、それほど主要な研究対象とならず、もっぱら適用範囲の確認や、移転可能性の問題など実務への適用研究に力が注がれて来たことが指摘できる。その結果として、セミナー開催後、比較的短期間で交通機関選択などに実務的成果を残すことができたと考えられる。同時期、海外では、1980 年には非集計分析の総括を行った国際会議 (International Conference on Research and Applications of Disaggregate Travel Demand

Models) がイギリスの University of Leeds で開かれ、非集計分析に対して当初の期待の大きさの反動として様々な批判もあった (Ortuzar and Kirby (1982))。それに呼応するように、1981 年にオックスフォードで開かれた会議では、交通の分析対象として、活動を中心に据えたアクティビティ分析が登場してきた (Carpenter and Jones (1983))。アクティビティ分析は交通行動理解のために非常に有効な視点を提示したが、これまでのところ強力な方法論がない (Kitamura (1988)) ため、実務ベースに用いられた例がほとんど存在せず、研究自体も減少気味であった。しかし、近年環境問題が深刻化し、交通の環境に与えるインパクトが世界規模の問題として取り上げられるようになり、その都市政策上への反映である Sustainable Development (持続可能な発展) に関連して、交通を適正な量に管理するための交通需要マネジメントの効果測定などには、活動との関連性に着目することが有効であるとの考え (北村 (1996)) から、アクティビティ分析は、この数年再び注目を集めるようになった。

一方、非集計行動モデルの理論が確立した直後の、ミクロ計量モデルの理論的研究を集めた Structural Analysis of Discrete Data with Econometric Applications (Manski and McFadden eds. (1981)) には、先に述べたアクティビティ分析以外にも、本論文で研究テーマとして取り上げる問題についての研究が見られる。例えば意思決定者の異質性についてのランダム係数モデルの研究 (Fischer and Nagin (1981)) やパネルデータを用いた動的分析 (Heckman (1981)) などである。その後、1980 年代前半からの、アメリカ合衆国において、時のレーガン政権の公共投資に対する予算削減政策を受けた連邦運輸省の交通研究に対する予算の削減や、理論的にそれほど重要な新展開がなかったことなどが原因で、非集計分析に対する研究にそれほど大きな成果は見られない。しかし、1980 年代後半から、いくつかの興味深い理論的新展開があり、再び非集計分析の研究が盛んになり、1991 年にカナダで開かれた国際交通行動学会 (ICTB; 現 IATBR) では、仮想の状況に対する意思表示である S P (Stated Preference) データを用いた分析など、これまであまり用いられなかったデータソースを用いた分析手法の理論的發展などがあり、再び多くの理論的研究成果を見ることができるようになった。現在ではこの非集計モデルの理論的發展の流れの中で、複数データセットを用いたモデルの研究や誤差分布の精緻化の研究などが盛んになりつつある。このような非集計モデルの理論的展開で、本研究は当初から非集計モデルの問題点とされていた、効用関数から抜け落ちた変数 (Omitted Variable) によるバイアスの問題を 2 章で取り扱い、3 章で S P データを採用して個人の効用パラメータの異質性の



問題に焦点を当てて分析を行う。4章、5章ではパネルデータを用いて、個人の異質性を考慮したモデル分析と後述するパネルサンプルの消耗によるバイアスを考慮したモデル分析を行う。

このように概観できる非集計モデルの発展の歴史に対し、現時点での研究の方向性について考える。これまでに述べたように、より細かな計画に対応できるように、交通行動モデルの精緻化は、一つの重要な研究課題として今後もその重要度は高まって行くであろう。同様に、ミクロ的視点とマクロ的視点に基づいたモデルの融合も、集団内での他者との関係が、現実の交通選択に大きなウェイトを占めるようになっていくに従って、その必要度を増して行き、より重要な研究課題となって行くであろう。同じように、近年のモデル論の発展の中の大きなウェイトをしめる、S Pデータやパネルデータなどの古典的クロスセクショナルデータ以外のデータソースを用いた分析や、大規模クロスセクショナルデータと小規模パネルデータの統合など異種データの複合的分析は、非集計分析の対象拡大とも関連して今後ますます増えてくると思われる。これら非集計行動モデル発展の背景には、心理学やマーケティングリサーチで発展した統計的分析手法の援用によって、これまで効用関数という一種のブラックボックスとして扱われてきた交通行動の意思決定構造に、2章で述べる知覚値や態度といった潜在的な要因を明示的に取り込む手法の発展がある。このような分析の考え方や統計的手法の発展は、様々なデータがいったい何を測定し、それが他のデータとどのような関係にあるのかということを明確にすることを可能にした。

ここで、マーケティングリサーチの分野で一般的に用いられている、時間的経過を考慮しない単一時間断面での消費者意思決定構造を図1-1に示す。この中で、選好の意思表示（S P）と態度の指標および知覚値の指標の3つが交通行動分析に主に用いられる意識データである。このような意識データについては、非集計行動モデルの創成期から、交通行動理解のために重要であることが指摘され、1970年代にこれらを取り込んだ非集計行動モデルの研究が盛んになった（Pas（1990））が、非集計行動モデルの背景に仮定される効用最大化原理という明快な理論と整合性を保てる分析手法が確立していなかったことなどによって、1970年代後半には下火になり、その後定性的な意識分析が、交通行動を理解するために行われてきた。しかし、マーケティングリサーチの分野ではその重要性がより強く認識されており、これらのデータを用いた定量的な分析手法の発展がなされてきた。これらの意識データの中で、選好意識を

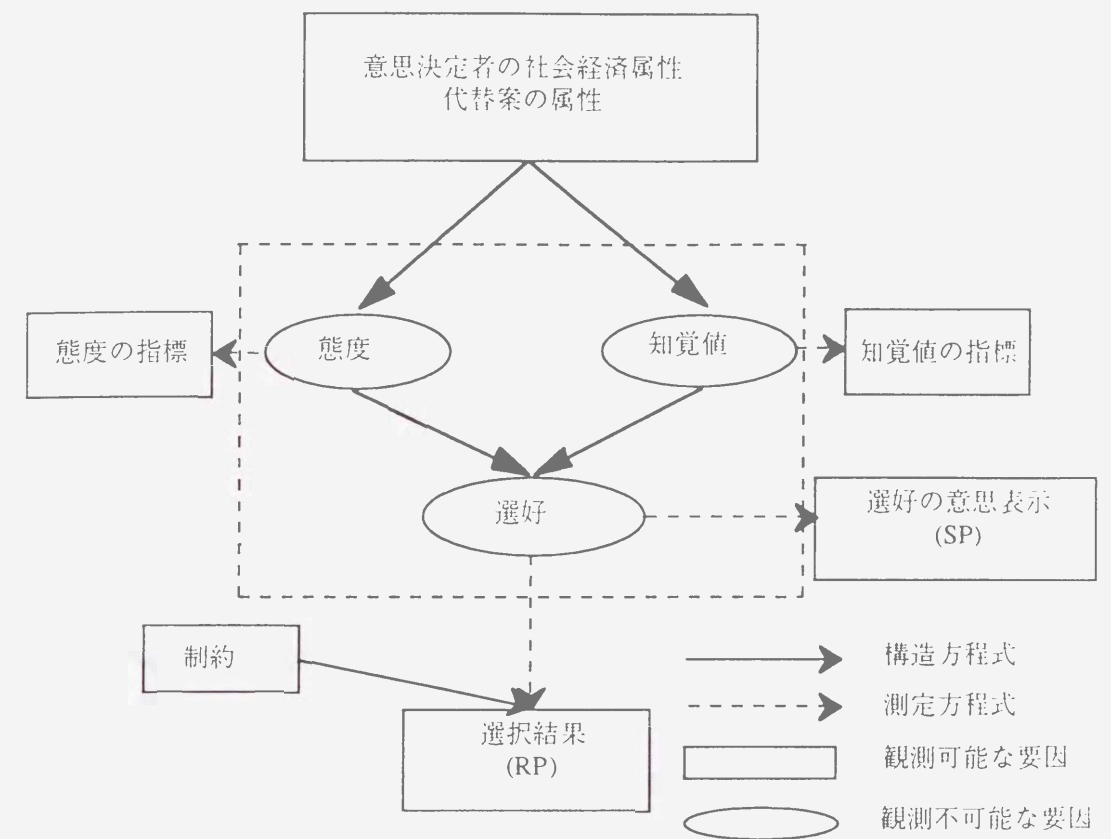


図1-1 単一時間断面での消費者意思決定過程解析のフレームワーク （Morikawa（1989））

表すデータは、被験者に仮想の選択状況を提示し、選好を尋ねたときの回答を用いることが多く、一般にS P（Stated Preference）データと呼ばれている。そして質問の形式により選択データ、ランキングデータ、評点データ、トランスファープライス<sup>1</sup>などのマッチングデータ<sup>2</sup>が存在する（森川（1991））。その他の意識データは測定しようとする対象により、知覚値指標と態度指標の2種類に大きく分けられている。計量心理学の分野では、S Pデータについては、属性の組み合わせをいくつか提示した中での選好を分析するコンジョイントアナリシスなどの手法が注目され、知覚値・態度データについては、潜在的な意識構造を明示的にモデル化し、これらのデータを定量的に分析する手法がそれぞれ発展してきた。知覚値・態度データの特性とその分析手法については2章及び3章で、S Pデータの特性とその分析手法については3章で詳し

<sup>1</sup> 転換価格。現在の利用手段と代替案が選好無差別になる費用・料金。

<sup>2</sup> より望ましい選択肢と選好無差別となるように、代替案の特定の属性値を答えてもらうもの。費用だけについて答えたものがトランスファープライス。

く述べる。

同じように、新たに非集計分析に用いられるようになってきた多時点データは、その収集方法や対象とするサンプルなどによりパネルデータ、多時点クロスセクションデータ、集計時系列データなどに分けられる。このうちのパネルデータについては4、5章で用いその特質などについてはそれぞれの章で詳しく述べる。先に述べたアクティビティアナリシスとの関連から、個人のトリップの記録を日記形式にしたトリップダイアリーデータ、活動も含めて一日もしくは数日間の日誌形式で採取するアクティビティダイアリーデータなどのダイアリー形式のデータも、非集計分析のフレームで研究される例がその大半を占めている。これらのデータの収集方法も、従来の調査票形式の家庭訪問調査や郵送調査だけでなく、携帯型パソコンを用いて、被験者の応答によってその場で質問を変動させることが可能な応答型調査や、マーケティングの分野では、オンライン化されたバーコード読取装置などを用いて物品購入履歴のデータを即時集計が可能な、スキャンパネルデータなどを用いることもすでに始められている。交通行動の分野では、このようなスキャンデータに近いものとして、警察が通称Nシステムという自動車ナンバー自動読み取り装置を各地の高速道路や大都市周辺の幹線道路に設置しているが、プライバシーの保護の問題などから、盗難車の捜査に利用されている程度であって、自動車の動きの把握に使われる予定はない。同じように自動車のナンバーをスキャンする、通称Tシステムという旅行時間情報システムも大都市や観光地周辺で稼働中であるが、これも現在のところ、所要時間や渋滞情報などの旅行時間情報の提供に用いられている程度であって、大規模に自動車の移動を把握する計画は存在しない。これらのナンバー読み取り装置はいずれも非常に高価であって、これらのシステムを用いて車の移動の把握するよりは、近年国内で普及のめざましい、カーナビゲーションシステムを用いて、地図上での移動を逐次保存することによって、対象となる車の移動を調査する方が現実的であると言える。

活動を含めたダイアリー形式のデータは、これまで研究レベルで収集はされているが、これまでのところ、交通と活動を含めた包括的な非集計モデルの分析フレームは存在せず、その一部の分析に非集計モデルが用いられることが多い。その中にはトリップチェーン単位での効用最大化や、時空間プリズム内での効用最大化を仮定するモデルなどが存在する。ただし、交通需要を異地点活動の派生的需要という単方向的な見方に基づいたアプローチでなく、交通と活

動・情報交換の相互関係をダイナミックにモデル化する必要が、今後21世紀に向けて大きく変動が予想される情報・通信技術の展開、超高齢化社会の到来、新しいコンセプトに基づいた移動手段の登場、といった状況下での需要予測には必要となる可能性が高いと言える。非集計離散型選択モデルを用いた分析の対象として、都市圏での日常的交通機関選択や立地分析など多方面に適用例が見られると述べたが、適用領域も交通計画の必要な範囲が広がるにつれて更なる広がりを見せている。例えば、航空需要の拡大に対応した航空計画のための国際航空需要の分析や、通勤・通学などの日常交通だけでなく、買物・観光などの非日常交通などへの適用、交通選択行動だけでなく、時間利用などの活動選択モデルなどへの適用が多く見られるようになってきた。

選択ルールの特定化についても、今後研究の発展して行く分野と考えられる。主に交通機関選択での適用例が多い非集計行動モデルは、所要時間、費用などの選択肢や意思決定者の属性を用いた加法型の効用関数を仮定し、最大効用をもたらす代替案を確率的に選択すると仮定する事が多い。このような場合に注意すべき点として、効用最大化行動を仮定して構築したモデルシステムは、効用最大化という仮定そのものは検証できないことが指摘され、一般に「モデルの循環」と呼ばれている（Lave and March（1975））。また、加法型の効用関数を用いることは、効用関数に含まれる総ての属性を同列に並べて評価し、選択しているという仮定も存在している。この場合、同一の効用レベルを達成する、つまり無差別曲線上に存在する属性の組み合わせは、無限に存在することになる。このことは各属性間で補償が可能であることを示しているため、一般に補償型の選択ルールと呼ばれている。交通行動の選択ルールが、常にこのような補償型であるとは限らない。実際問題として、観光目的地の選択を例にとりて考えてみると、観光目的地は選択肢が非常に多いことが一般的であり、このような場合、意思決定者は総ての選択肢を同列に並べ、その属性を同時に比較しているとは考えにくい。それよりは、最も重要な属性に着目し、比較優位である選択肢を階層的に絞り込んで行く辞書編纂型ルールや、各属性に最低水準を設定し、それを総て満たしている選択肢に絞り込んでいく連結型ルールなど、属性間での補償を許さない非補償型の選択ルールに基づいていると考えた方が妥当性が高い場合も多く存在する。しかし、非補償型選択モデルの方が行動記述の面で優れていると考えられる局面において、このような選択ルールに対応した計量モデルもいくつか提案されているにもかかわらず、非補償型ルールに基づいた選択モデルは、これまで用いられることは少な

った。これは、 加法型の効用関数は効用パラメータの与え方などにより、非補償型選択ルール  
の近似を与えることが可能であることと、非補償型選択ルールに基づいた計量モデルは、ロジ  
ットモデルのようなシンプルな形で表現されているものが無く、推定作業や予測計算に煩雑さ  
が増えることが多いなどの理由によると思われる。しかし、計量的な需要予測だけでなく、公  
共交通の需要喚起や、都市中心部の活性化などのマーケティング戦略作成なども土木計画の一  
環として扱われるようになり、細やかな計画立案の必要性が高まってきた。このような場合は、  
非補償型の選択ルールに基づいた分析の有効性が高く、補償型の選択モデルで近似することが  
困難であるため、近年再び非補償型選択モデルの再検討が行われている。

また、これに関連したケースとして、個人ごとの選択肢集合の正確な把握の必要な局面とい  
うものも増えている。例えば、普段バスを利用しない層に対して、バスを含めた公共交通機関  
の需要喚起策を策定する場合、選択肢集合形成過程を考慮していない交通機関選択モデルでは  
不十分な場合がある。なぜなら、このような層は、もともとバス路線の存在を知らず、バスが  
選択肢集合に入っていない人も含まれると考えられ、バスが意思決定者全員の選択肢集合に、  
確定的に含まれたモデル、つまり、全員がバスの路線を正確に把握しているという仮定の下で  
の需要予測は、バスがもともと選択肢集合に入っていない層の行動モデルのパラメータにバイ  
アスを与え、バスのシェアを過大に推計することになる。このようなバイアスを回避するため  
には、選択肢集合形成過程を含めたモデル化が有効であり、需要喚起策は選択肢集合形成段階  
に働きかけるものが有効になると考えられる。このような選択肢集合形成過程は、先に述べた  
ように、非補償型ルールに基づいていると考えることが妥当な場合が多く、この意味でも、非  
補償型選択ルールを含めたモデル構築の必要性は高いと言える。

モデル論的見地からとは別に、一般的・社会的背景から今後交通行動モデルに必要となってい  
くるものとして、環境・エネルギー問題に対する交通からの取り組みをどのように評価・分析  
していくのが重要になって行くであろう。環境・エネルギー問題は、非集計行動モデルが仮  
定する個人レベルでの効用最大化の結果として、長期的・社会的な目標が達成されることが難  
しい問題が多く、その目標達成のための交通需要マネジメント策には、個人の欲求を抑える方  
向に働く政策が増えてくるであろう。このような状況下での交通行動分析には、各意思決定者  
を、同じ集団に属する他人とは独立に意思決定を行う主体として取り扱うのではなく、流行、

同調・非同調といった他者との相互関係を含めた行動分析や、社会的合意形成過程のモデル化、  
社会的判断、説得と誘導などの組織論・対人行動論などをふまえた分析が必要になってくると  
考えられる。これらはミクロ的視点に立つ分析とマクロ的な視点での分析を一体化する方法論  
を必要とし、現在でも非集計モデル適用の際の需要の集計化の問題とあわせ発展の方向として  
提示されると思われる。

### 1-3 本論文の構成

本論文は、以下のような章立てで構成される。まず1章で研究の背景と目的について述べ、  
交通需要予測のこれまでの研究を、非集計分析という切り口で概観し、変容する社会環境との  
関連から、今後の研究発展の方向について展望する。続いて2章で主観的潜在変数を効用関  
数に導入した交通機関選択分析を行う。主観的潜在変数の同定には線形構造方程式モデルを用  
いるが、潜在変数を含むモデル一般の特性について検討し、線形構造方程式モデルと離散型選  
択モデルの統合をベースに研究を進め、その事例研究により得られた結果から本研究で提案し  
たモデルの有効性の検討を行う。3章では個人の効用パラメータの異質性を非集計分析に導入  
する手法の検討を行う。本章で行う研究の特徴としては、個人の効用パラメータの異質性が現  
れやすいとされるS Pデータを用いることである。特に同一個人から選好の情報を繰り返し得  
ることが容易であるというS Pデータの特性を生かし、同一個人から繰り返し複数回の選好情  
報を得たS Pデータを用いて、個人の効用パラメータの異質性を考慮したモデルを構築する。  
まず、既存のトリップ目的や社会経済属性など、事前に外生的な基準を定め、それに沿ってセ  
グメント分けを行うアプリオリセグメントの問題点を提示し、その対策として意思決定者ごと  
にモデルを推定し、得られた個人パラメータによるアプリオリセグメントや、アンケートによ  
って得られた態度指標に線形構造方程式モデルを適用し、得られた潜在的態度変数の推計値に  
よるアプリオリセグメント手法の提案を行いその事例研究を行う。また、非集計行動モデルへ  
の潜在セグメントモデルの適用を提案し、個人属性や態度変数を用いた事例研究を通じて、潜  
在セグメントモデルの有効性を検証する。4章、5章ではパネルデータを用いた動的モデルの  
開発を行う。まず、パネルデータの問題点と利点を列挙し、4章では、3章で用いた複数S P  
データと同じく、同一個人から複数の観測を行うというパネルデータの特長を用いて、潜在ク  
ラス分析の事例研究を行い、その有効性を再び検証する。このとき、時系列的に安定した非観

測要因による個人間の誤差相関を考慮したモデルもあわせて推定を行う。続く5章では、パネルアトリション（消耗）を考慮した選択モデルの構築を行う。パネルアトリションは、4章でモデリングを行った、個人の非観測要因に深く関連しており、この影響を考慮したモデル構築を行う。このとき、配布された調査票に対し、返答無しや記入漏れなどによる有効回答の減少を事前消耗と定義し、この問題とパネル調査での消耗現象を、同一の行動原理に基づくと仮定した分析フレームの中でモデリングし、事前消耗とパネル消耗の両方によるバイアスを同時に解消する手法を提案し、事例研究を行う。そして6章で本研究をまとめ、本研究で得られた知見と今後の課題について述べることとする。

## 第1章 参考文献

北村隆一(1996) 交通需要予測の課題：次世代手法の構築に向けて、土木学会論文集, No.530/IV-30, pp.17-30.

Carpenter, S. M. and Jones, P. M. (eds.) (1983) Recent Advances in Travel demand Analysis, Gower, Farnborough, England.

Fischer, G. W. and Nagin, D. (1981) Random versus Fixed Coefficient Quantal Choice Models, In Structural Analysis of Discrete Data with Econometric Applications, C. Manski and D. McFadden eds., M.I.T. Press, Cambridge, U.S.A., pp.273-304.

Heckman, J. J. (1981) Statistical Models for Discrete Panel Data, In Structural Analysis of Discrete Data with Econometric Applications, C. Manski and D. McFadden eds., MIT Press, Cambridge, U.S.A., pp.114-178.

Kitamura, R. (1988) An Evaluation of Activity-Based Travel Analysis, Transpotation, 15, pp.9-34.

Lave, C. A. and March, J. G. (1975) An Introduction to Models in the Social Science, Harper and Row, New York, U.S.A..

Manski, C. and McFadden, D. (eds.) (1981) Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, U.S.A..

Morikawa, T. (1989) Incorporating Stated Preference Data in Travel Demand Analysis, ph.D. Dissertation, Department of Civil Engineering, M.I.T.

Ortuzar, J. D. and Kirby, H. (1982) Fowarded, Transportation Research, Vol.16(A), pp.337-338.

Pas, E. (1990) Is Travel Demand Analysis and Modeling in the Doldrums?, Jones, P. (ed.), Developments in Dynamic and Activity-Based Approaches to Travel Analysis, Avebury, pp.3-27.



## 第2章 主観的意識要因を考慮した交通行動分析

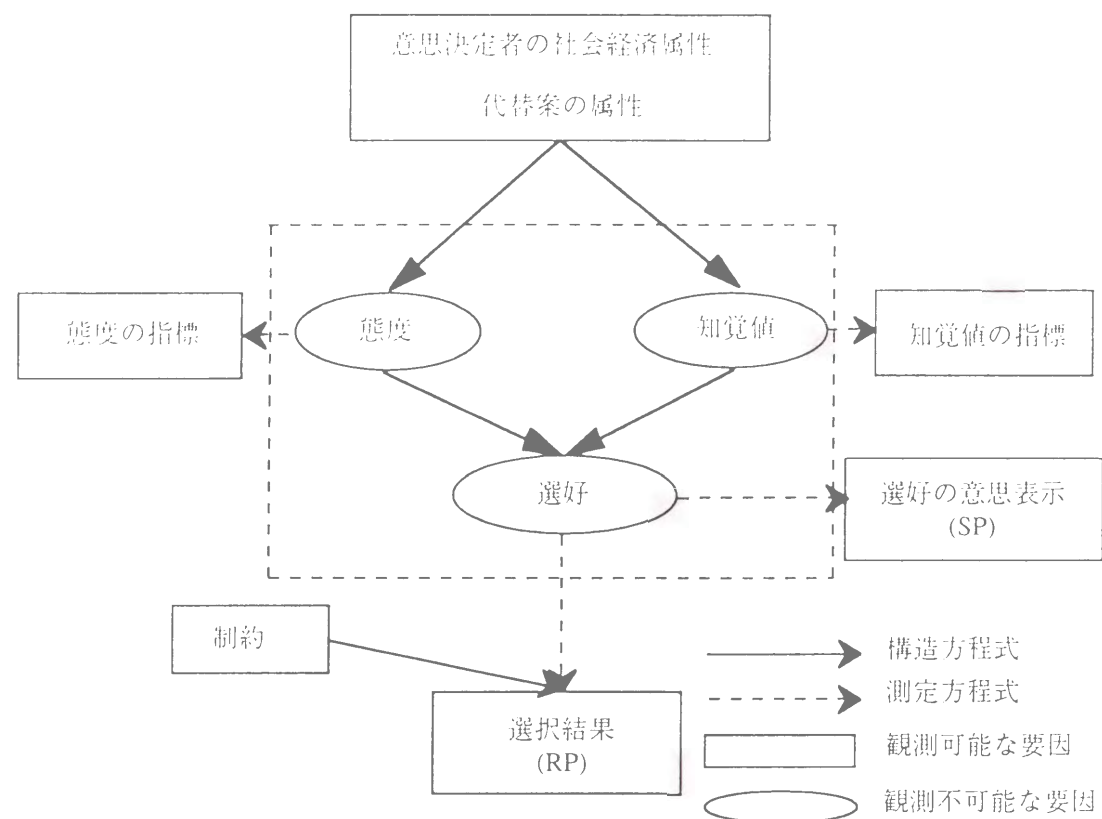
### 2-1 はじめに

交通需要予測における交通機関選択分析は、近年研究が最も盛んに行われている分野の一つである。特に、計量経済学で理論的發展が進められてきた離散型選択分析（Discrete Choice Analysis）を用いた、非集計行動モデルの適用によるところが大きい（Ben-Akiva and Lerman（1985））。非集計行動モデルの特徴としては、背景となる人間の行動理論が明快で、個人の選択行動メカニズムを表現できるため、政策の変化に対する予測に使いやすいことや、故少ないサンプルのデータで、信頼性の高いモデルが推定できることなどが挙げられる（森地（1984））。しかし、これまで適用されてきたモデルでは、選択行動を説明する要因として、所要時間・費用などの定量化が可能なものしか用いられないことが多かったが、実際の選択行動メカニズムは、定量化可能な要因以外にも、潜在的な主観的要因によって規定されていることは自明であり、また指摘・研究が非集計モデルの創成期から現在まで続いている（例えば Recker and Golob（1976）、鈴木ら（1986））。そこで、本研究は、潜在的な要因を取り入れた交通行動分析の方法論を展開し、それに対する実証的研究を行うことを目的としている。

本章で提案する方法の概略は、アンケート調査によって得られた主観的心理データを用いて、潜在的な要因を取り入れた交通機関選択モデルを構築するものである。この考え方の背景にあるのは、ミクロ経済学の基礎である消費者行動理論が、マーケティング・リサーチの分野で見直されていることにある。これまでのミクロ経済学では、ブラックボックスと見なされてきた消費者の行動機構を、潜在的（Latent）な心理的要因を明示的に取り入れることによって、その機構を明らかにしていこうとするものである。この新しい消費者行動の構造を、パス・ダイアグラムで表したものが1章でも示した図1-1である。この図において、楕円は潜在的な要因であることを示し、長方形は観測可能な要因であることを示している。

代替案と意思決定者の属性、選択結果という2つの観測可能な変数は、態度、知覚値、選好という3つの潜在的な要因によって結びつけられている。知覚値とは、意思決定者が選択肢の属性を認識している値で、その人の社会経済的属性と、その人が得ている代替案の属性に関する情報に影響されていると考えられる。態度とは、意思決定者が代替案の属性や、代替案そのものに対して抱いている感情で、意思決定者の社会経済属性や文化的背景に影響されていると考えられている。選好とは、選択行動を決定するための潜在的な変数で、ミクロ経済学では「効用」





という指標で表されることが多い。

そこで本研究は、潜在的要因を他の定量化可能な要因と整合性が取れる変数として表すため、定量化可能な要因と心理指標の関係を、潜在変数を含む方程式の集合としてモデル化する方法

## 2-2 主観的潜在要因と数理心理学データの特性

いわゆる Attitude-Behavior モデルに代表されるように、交通行動分析に潜在要因を取り込むための研究は非集計モデルの創成期から続いている。Lovelock (1975) は意思決定過程をフローチャートで表し、潜在的な態度や知覚値またそのときの状況制約が、交通行動を把握するために重要だと述べている。同じく Dobson (1975) は交通行動に関連する意思決定原理と、政策変数の変化に対する行動変化の評価の方法を議論する中で、潜在変数の重要性に着目し、そのデータ収集などに関して知見を述べている。Recker and Golob (1976) は通勤交通の機関選択分析で、交通制約に着目してサンプルを分類し、因子分析を用いて交通機関選択に重要な因子を各グループごとに抽出し、その因子を代表する指標を用いてロジットモデルを推定している。Dobson and Tisher (1977) は知覚値と選好の交通行動との関連について研究を行なっている。ここではデータをアンケートによる意識データ、ネットワークデータによる選択肢の客観的変数、意思決定者の社会経済属性の 3 種類に分け、それぞれが交通行動に与える影響とデータ相互の関係を調べ、知覚値が交通機関選択に重要だと結論づけている。

Koppelman and Pas（1980）は選択肢の客観的属性と知覚値、知覚値と選好、選好と選択の関係をモデルで表し、因子分析を用いて知覚値と態度（feeling）を求め、ロジットモデルを用いて知覚値だけの選択モデルと知覚値と態度を用いたモデルの比較を行なっている。Koppelman and Lyon（1981）は2種類のデータを用いて Koppelman and Pas（1980）と同様の方法で選択肢別の態度を用いたモデルと選択肢間で共通の態度を用いたモデルの比較を行なっている。Morikawa et al.（1990）は本研究のベースとなるもので、消費者行動仮説に基づき、SP データや意識データを統合した交通行動モデル構築のサブモデルとして、線形構造方程式モデルを用いて潜在変数を同定する方法を提案し、知覚値を取り入れた交通機関選択分析を行なっている。一方国内では、非集計モデルが導入された当初から潜在変数の重要性は指摘され（太田、原田（1982））、河上・広島（1985）は、自動車と鉄道の交通手段転換モデルに、主観的評価値を取り入れている。ここで主観的評価値とは、アンケートで得られた鉄道及び自動車のサービスレベルに対する総合的満足度（「非常に満足」から「非常に不満」まで7段階）のデータを、心理尺度構成法の1つであるカテゴリー判断の法則を用いて、客観的要因から計算されうる連続変数に変換したもので、それを効用関数の説明変数に用いている。鈴木ら（1986）はデータを、被験者に所要時間などを尋ねたもので、意思決定者が知覚している値をそのまま用いることが可能な回答値、主観的に選択肢を評価した値である評価値、客観的判断基準によって定められた設定値の3種類に分類し、それぞれのデータによるモデルの比較を行なった結果、回答値を用いたモデルと評価値データを設定値に追加して用いたモデルは、従来の設定値のみによるモデルと比較して適合度が高いと結論づけている。渡辺ら（1987）は新交通システム導入の際の需要予測に、選択不選択理由の分析およびサービス要因の満足度評価の分析を行ない、新交通システム選択の際に重要になる潜在要因を定め、そこで得られたサービス要因と利用意向データの関係を数量化 II 類を用いて分析し、それらの要因を代表する指標を用いて選択モデルを推定している。青山ら（1988）は交通機関選択の重要な要因として経済性、高速性、快適性を取り上げ、それらを表す指標として所要時間、交通費用、疲労を提案している。ここで疲労とは、潜在要因である快適性を定量化するために、個人のエネルギー消費量を表す式を提案し、それを用いた交通機関選択モデルを推定している。Balace（1988）はマニラの都市交通に関するアンケート調査で得られた心理的指標から、交通機関選択に重要な要因を因子分析を用いて分析し、その指標を取り入れた効用関数を推定している。上記の問題点を克服するために、Morikawa（1989）は、構造方程式モデル（LISERL モデル）を用いて、潜在変数を定量化する方法論を展開した。

本研究は、この研究で提案されたモデルを基礎として、その推定法の改善および実証的研究に主眼を置くものである。

## 2-2-2 交通機関選択に影響する潜在要因

2-2-1 で述べたように、従来の交通機関選択モデルには、旅行時間や費用といった定量化が容易な要因だけを取り入れているものがほとんどであった。しかし、多くの研究者が指摘しているように、実際の行動は、様々な定性的要因が影響を与えていると考えられる（例えば Ben-Akiva and Morikawa（1990））。特に本論文の事例研究では、鉄道と自動車の選択を取り扱うので、具体例としてこの2つのモードに関係する定性的要因を挙げた。

### （1）信頼性（Reliability）

ここで例えば「到着時刻の信頼性」を定量的に表現すると、「到着予定時刻が許容時間内に収まる確率」である。鉄道は、専用軌道を持ち、ダイヤグラムにしたがって計画的に運行されているが、自動車は他の交通手段が混在する一般道路を走るため信頼性は低いと思われる。しかし、鉄道も乗り換えの回数や、アクセス、イグレスの交通機関によって信頼性は変化するとと思われる。

### （2）情報利便性（Information Availability）

利用者が、必要としている情報を得ることの難易度。鉄道ならば、時刻表ではほぼすべてのことがわかるが、自動車の場合、地図で経路選択は可能であるが、所要時間などの情報は得ることが困難である。鉄道と路線バスの情報利便性を比較してみると、バスはその存在さえ知られていないこともあり、路線のわかりやすさなどでも鉄道の方が大きく上回っていると思われる。

### （3）快適性（Comfort）

乗り物の「乗り心地」、「居住性」を表す。通勤電車では、常に着席可能ではなく、空調も自由にはならず、混雑していることが多い。特に地下鉄は、車窓の景色・騒音の点からも快適性は低いと思われる。一方、自動車は必ず着席でき、空調も自由になり、音楽を聴くことも可能である。また、運転者か同乗者かで、快適性は異なると思われる。しかし、長距離旅行の場合は、一般に鉄道の方が座席もゆったりしており、車内で飲食のサービスがあり、長距離のドライブより快適性は高いのではないかとと思われる。

#### （４）安全性（Safety from Accident）

その交通機関の、事故の起こりやすさ、また事故が起きたとき受ける被害の大きさを示すものである。鉄道は、専用軌道を走行し安全装置も備わっているが、一度事故が起こるとその被害は大きい。自動車は、様々な交通手段と道路を共用しており、走行は運転者の判断に依存しているため、安全性は低いと考えられる。

#### （５）防犯性（Security）

その交通機関利用時に、犯罪に巻き込まれる可能性である。鉄道は、車内においても駅においても乗客の数に対し乗務員数は少ないので、犯罪に巻き込まれる可能性は大きい。わが国において交通機関選択の際に防犯性はあまり問題にならないが、海外の大都市において防犯性は非常に重要な要因である。

#### （６）利便性（Availability）

その交通機関の利用しやすさである。鉄道を利用する場合は、必ず駅まで行かねばならず、かつ駅でしか降りられず、運行時刻もあらかじめ決められているが、自動車は、戸口から戸口への移動が可能で、出発時刻も自由に決めることが可能である。しかし、わが国の都心部への旅行では、駐車場の不足から、かえって自動車の方が利便性が低いこともある。

#### （７）イメージ（System image）

利用者が、その交通機関に対して抱いている心象であり、これは人によって、鉄道に対しても自動車に対しても異なる

#### （８）プライバシー（Privacy）

利用者のプライバシーが、どの程度守られるかを示す。鉄道は、多くの見知らぬ人が1つの車両内に存在するため、自動車の方が圧倒的に優れていると考えられる。

### 2-2-3 認知的不協和とカテゴリカルデータの特性

認知的不協和とは自分の行った行動が、自己の評価基準に基づいて評価を行った場合最良の行動でない場合におこる。一般的に意思決定者はそのような不協和状態に対し、認知的不協和の存在を解消する方向で、評価基準や評価項目の変更などを行うことによって再評価を行う。つ

まり、自己の行動の正当化の一種であると考えられる。これは一般に消費者の購買行動に現れる（Assael（1987））。商品の情報収集を行いブランド間に違いが見られないとき、購入した後に購入したブランドを支持するような情報収集を行い、その商品の全体的な評価を形成するという過程をたどる（秋山（1995））。本研究で扱う交通機関選択においては、認知的不協和とは自己の選択した代替案に対して、後にその選択を評価したとき、それがベストの選択という自信がもてない状態が認知的不協和の状態で、意思決定者はそのような状況を解消するために自己の選択した代替案に有利な再評価を行う。本研究は、RP（Revealed Preference）データ、つまり被験者の実際の選択結果が明らかな行動に対して取られたデータであるため、先に述べた認知的不協和の状態になっている場合、アンケートの中で代替案の主観的評価を尋ねた項目に、自分の選択した結果とアンケートの回答が矛盾しないよう選択したモードを過大に評価する可能性がある。もしそのような場合、アンケート調査より得られた主観的評価値は、図1-1で示した構造に基づいて主観的評価を測定したことにはならず、実際の選択結果とアンケート調査の時間的ずれの影響を含んだものになっている。このときモデルのパラメータの推定値には、Omitted Variable によるバイアスが存在する恐れがある。この時間的ずれを含んだ意思決定者の

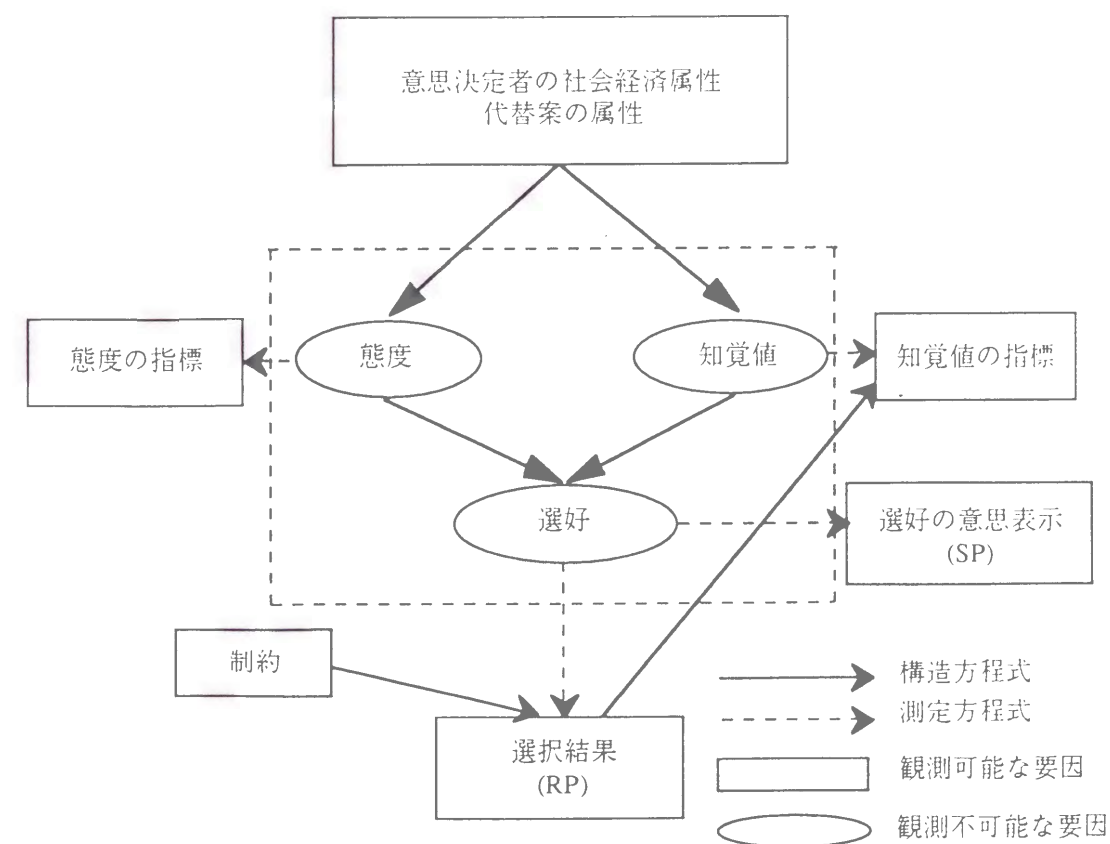


図2-1 認知的不共和を考慮した意思決定過程のパスダイアグラム

意識構造をパスダイヤグラムで示したのが図2－1である。このパスダイヤグラムでは、知覚値指標に選択結果からのびる矢印が、認知的不協和解消行動を示している。

知覚や態度を測定するには尺度を与えて測定を行う。通常尺度は一定の単位があるかどうかによって、〔比例尺度、間隔尺度〕か〔順位尺度、名義尺度〕にわけられる。比例尺度、間隔尺度で測られたものは数量データであり、順位尺度、名義尺度で測られたものは定性的、質的データである。質的変数は、カテゴリカル変数、データとして扱う場合にはカテゴリカルデータと呼ばれる（柳井ら（1990））。カテゴリカルデータにはある交通機関を選んだかどうかや交通が発生したかどうかのようなある事象が生じたかどうかのみが記録される場合のようにカテゴリの数が2つの場合と、態度測定や知覚値測定、社会調査の項目に対する反応や、多肢選択の質問に対する回答のようにカテゴリの数が3つ以上の場合がある。前者は2値変数（dichotomous variables）、後者は多値変数（polytomous variables）と呼ばれる。平均、分散やこれらに基づく回帰分析などの多変量解析手法は本来的に計量データに適用が可能なものであって、質的データにはそのままでは適用が困難である。通常、知覚値指標や態度指標などのデータは、満足度の評点づけや質的データとして得られる場合が多い。それは評点付けデータの場合で数値化されているとしても、本来的に連続的な評点が離散的変数に変換されており、その間隔が等間隔であるという保証もない。5段階評価や10段階評価の場合には、それぞれ各個人の評点が5種類、10種類のカテゴリに反応した質的データと考えられる。先に述べたように態度測定などを計算機で処理する場合、“好き－嫌い”などの程度を表す言葉に便宜的に“1～5”などの数値を割り振ることが多い。これはあくまで便宜的なものであって、非常に好き（5）と非常に嫌い（1）の違いは5倍であるという保証はない。便宜的に割り当てた数値が例えば「好み」という変数を測定したものであり、その増加に対して近似を与えるとはいえ、これらのデータの平均、分散などの処理された数値をそのまま無批判的に用いることは、データの解釈や分析結果に対し誤った判断をする危険性がある。このような問題に対しこれらのカテゴリに意味のある数値を当てはめるための手法として数量化理論（林、駒沢（1982））が開発され、対数線形モデルなどの応用（松田（1988））などが考案されている。

2－3 モデルの定式化

2－3－1 潜在変数モデル（柳井ら（1990））

本研究では、交通機関選択において重要な役割を果たす潜在的要因を、定量化することを目的としているため、潜在変数を含むことができるモデルを考えねばならない。ここで扱われる潜在変数は、個々の観測個体の特徴を示すものであり、通常の統計モデルとは異なる性格をもつ。潜在変数を含むモデルを潜在変数モデルと呼ぶことにする。

潜在変数モデルには色々な種類があり、これを1．顕在変数が量的尺度（間隔尺度または比尺度）か、質的変数（名義尺度または順位尺度）であるか、2．潜在変数が量的変数か質的変数であるか、3．母数の間に構造をあらかじめ仮定するかどうか、という3つの基準によって分類する。

本章の分析では、顕在変数と潜在変数がともに量的変数であると見なし、事前の知識によって母数に特定の構造を仮定することにより、LISREL モデルを用いて分析する。

表2－1 潜在変数モデルの分類

| 顕在変数 | 潜在変数 | 特定の構造 | 分析法            |
|------|------|-------|----------------|
| 量    | 量    | 無     | 探索的因子分析        |
| 量    | 量    | 有     | 確認的因子分析、LISREL |
| 量    | 質    | 無     | 潜在プロフィール分析     |
| 量    | 質    | 有     |                |
| 質    | 量    | 無     | カテゴリカルデータの因子分析 |
| 質    | 量    | 有     | 項目反応理論         |
| 質    | 質    | 無     | 潜在クラス分析        |
| 質    | 質    | 有     |                |

2－3－2 LISREL モデル（Joreskog and Sorbom（1984））

LISREL モデルには多変量解析の手法としての2つの側面がある。1つは確認的因子分析モデルであり、もう一つはパス解析である。確認的因子分析とは、ある現象を因子分析する際には



予備知識を有していることが普通であるが、その場合にそれらの予備知識を仮説として分析モデルに組み入れ、その仮説の是非を確認する方法である。通常次の3つの仮定がおかれる。

- 1) 共通因子間の相関について特定の構造を仮定する。
- 2) ある変数は幾つかの因子とのみ関係し、他の因子とは無関係であることを仮定する。
- 3) 特定の独自因子間に相関があることを仮定する。

一方パス解析とは、いくつかの変数間に何らかの方向性をもった因果モデルを設定し、各変数を結ぶパスの影響の強さを定量的に推定するもので、因果関係があると考えられる変数間に線形回帰式を想定する。すなわち、因果関係の数と同じ数だけの方程式ができ、データから推定される回帰係数によって因果関係の強さを推定する手法である。

LISREL モデルでは、これらをさらに一般化し、従属変数と独立変数間の因果関係を潜在変数を介在させて、その間の線形関係とするモデルである。つまり因子分析モデルによって観測可能な変数から、その規定要因となる因子をモデル化し、それと同時にこれら潜在因子間の因果関係を、構造方程式として把握することを試みるモデルである。これらのことから、LISREL モデルは、図2-1の例のような変数間の因果関係を表すパス・ダイアグラムをモデル化する際に有効な手法であり、本研究においては非常に有効であると考えられる。

モデルは基本的に「構造方程式 (structural equations)」と「測定方程式 (measurement equations)」という2種類の式から成り立っている。変数間の因果関係は構造方程式で表され、構造方程式中の変数はすべて潜在変数とみなされる。測定方程式は、観測可能な多くの「指標 (indicators)」と潜在変数との関係を表す。構造方程式中のいくつかの潜在変数が直接観測可能ならば、それらの変数を構造方程式中で完全に計量された潜在変数として扱い、測定方程式中で潜在変数およびその指標として扱うことによって定式化の一般化を計っている。そのような定義のもと、モデルを一般形で示すと次のようになる。

構造方程式

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta} \quad (2.1)$$

測定方程式

$$\mathbf{Y} = \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (2.2)$$

ただし、

$\boldsymbol{\eta}$  = 潜在変数ベクトル

$\mathbf{Y}$  = 指標ベクトル

$\mathbf{B}, \boldsymbol{\Lambda}$  = 未知パラメーター行列

$\boldsymbol{\zeta}, \boldsymbol{\varepsilon}$  = 多変量正規分布 (MVN) に従うランダム項ベクトル

構造方程式中では潜在変数はすべて内生変数になっているが、パラメーター行列の  $k$  行目をすべて 0 におくことによって  $\boldsymbol{\eta}$  の  $k$  番目の変数は外生変数、つまり上で説明した完全に計量された潜在変数にすることができる。この LISREL モデルは多くの多変量解析モデルの一般形となっており、構造方程式だけを取り出すと同時重回帰モデルとなり、測定方程式だけの場合は因子分析モデルと同型となる。

このモデルの推定法を以下に示す。 $\mathbf{Y}$  をサンプル平均からの偏差として測定すると  $\mathbf{Y}$  の共分散行列は、

$$\mathbf{E}[\mathbf{Y}\mathbf{Y}'] = \mathbf{E}[(\boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon})(\boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon})'] = \boldsymbol{\Lambda}\mathbf{E}[\boldsymbol{\eta}\boldsymbol{\eta}']\boldsymbol{\Lambda}' + \mathbf{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda} + \boldsymbol{\Theta} \quad (2.3)$$

となる。

ただし、

$$\boldsymbol{\Phi} = \mathbf{E}[\boldsymbol{\eta}\boldsymbol{\eta}] \quad (2.4)$$

$$\boldsymbol{\Theta} = \mathbf{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}] \quad (2.5)$$

$$\mathbf{E}[\boldsymbol{\varepsilon}\boldsymbol{\eta}] = 0 \quad (2.6)$$

また、構造方程式モデルは、

$$\boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\zeta} \quad (\mathbf{I} \text{ は単位行列}) \quad (2.7)$$

と変形できるから、

$$\boldsymbol{\Phi} = \mathbf{E}[\boldsymbol{\eta}\boldsymbol{\eta}] = \mathbf{E}[(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\zeta}\boldsymbol{\zeta}'(\mathbf{I} - \mathbf{B})^{-1}] = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B})^{-1'} \quad (2.8)$$

となる。ただし、

$$\boldsymbol{\Psi} = \mathbf{E}[\boldsymbol{\zeta}\boldsymbol{\zeta}'] \quad (2.9)$$

(2.8) 式を (2.3) 式に代入することによって、

$$\mathbf{E}[\mathbf{Y}\mathbf{Y}'] = \boldsymbol{\Lambda}(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B})^{-1'}\boldsymbol{\Lambda}' + \boldsymbol{\Theta} \quad (2.10)$$

で表される指標の理論的共分散行列を得る。これをサンプル共分散行列と適合させることによって、未知パラメーター行列  $\mathbf{B}, \boldsymbol{\Lambda}$  を推定することができる。一般には、すべての変数が正規分布しているという仮定のもとに最尤推定量を用いることが多い。LISREL モデルを推定するコン

ビュータ・プログラムとして LISREL (Scientific Software International)、LINCS (RJS Software) などがある。

### 2-3-3 フレームワーク

ここでは、2-1で概略を述べた消費者行動分析の考え方に基づいて、数理心理学データを交通需要予測モデルに適用する方法論を展開する。これまで述べてきたように態度や知覚値といった潜在変数は、意思決定者の交通機関選択に大きな役割を果たしていると考えられ、そのような潜在変数の指標である数理心理学データを、交通機関選択モデルに取り込むことが、モデルの交通行動分析への適用性を大きく改善すると考えられる。

本研究は、そのようなデータをモデルへ取り込む手法を提案するものである。図2-3は本研究での数理心理学データの役割と、潜在変数との関係を表したものである。この図で表されているように、観測可能な要因と潜在的な要因は、構造方程式で結ばれ、数理心理学的データは、

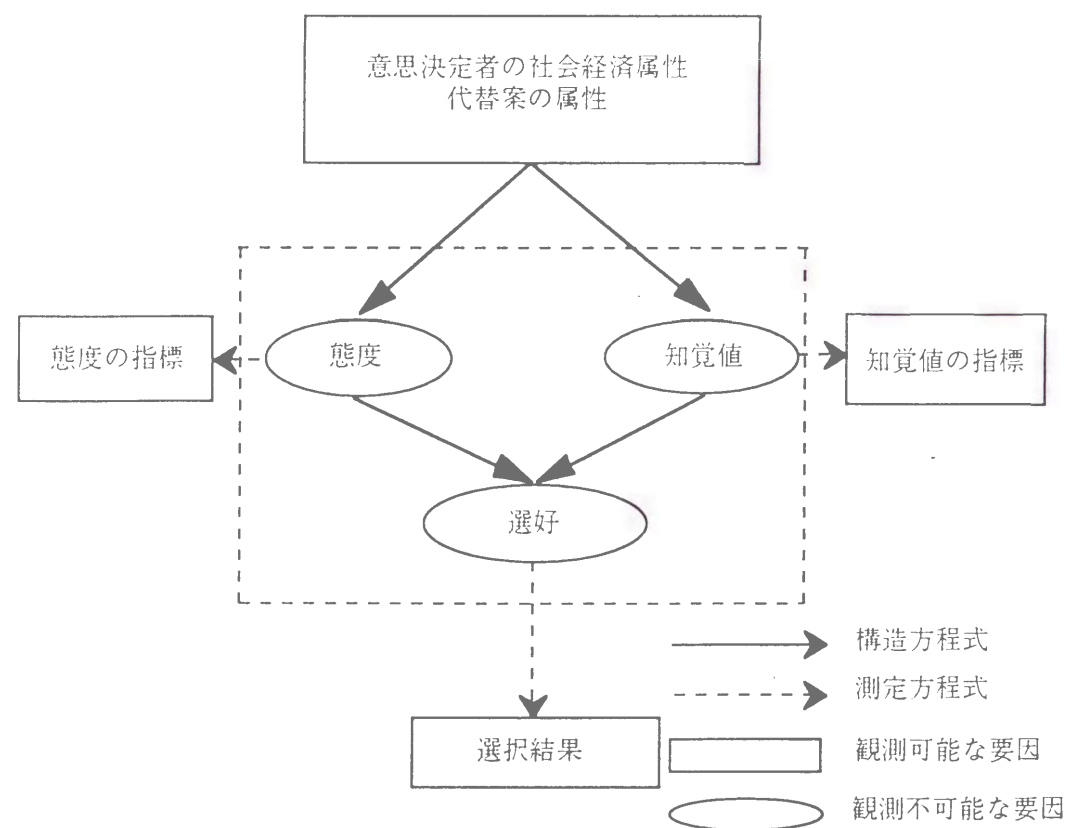


図2-3 本研究での数理心理データの役割

潜在的な要因と測定方程式で結ばれている。

以下、ランダム効用理論に基づく離散型選択モデルと、LISRELモデルを用いた本研究での潜在的な要因を含む離散型選択モデルの定式化を行なう。ロジット・モデルやプロビット・モデルに代表されるランダム効用に基づく離散型選択モデルはそれ自体、「効用」という潜在変数を含むモデルである。このとき、LISRELモデルの用語に従えば、効用関数が構造方程式に、選択を表すダミー変数が測定方程式に相当する。このシステムは次のようなフレーム・ワークで表すことができる。なお、以下の定式化では簡単のために二項選択モデルを例に説明し、変数はすべて2つの代替案の差で表されているものとする。また、直接に観測できない潜在変数はアスタリスク(\*)を付けて表している。

#### 構造方程式

$$u^* = \mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{x}^* + v \quad (2.11)$$

$$\mathbf{x}^* = \mathbf{B}\mathbf{s} + \zeta \quad (2.12)$$

ただし、

$u^*$ =選択モデルの効用

$\mathbf{x}$ =選択モデルにおける観測可能な(客観的な)説明変数のベクトル

$\mathbf{x}^*$ =選択モデルにおける潜在的・潜在的な説明変数のベクトル

$\mathbf{s}$ =構造方程式における $\mathbf{x}^*$ を形成する客観的説明変数のベクトル

$\mathbf{a}, \mathbf{c}, \mathbf{B}$ =未知パラメータの配列

$v \sim N(0, 1)$  に従う効用関数のランダム項

$\zeta \sim MVN(0, \Psi)$  に従うランダム項

#### 測定方程式

$$d = \begin{cases} 1 : \text{if } u^* \geq 0 \\ -1 : \text{if } u^* < 0 \end{cases} \quad (2.13)$$

$$\mathbf{Y} = \mathbf{\Lambda}\mathbf{x}^* + \varepsilon \quad (2.14)$$

ただし、

$\mathbf{Y}$ =アンケートで得られた主観的評価値ベクトル

$\mathbf{\Lambda}$ =未知パラメータ行列

$\varepsilon \sim MVN(0, \Theta)$  に従うランダム項

このシステムでは、(2.11)式と(2.13)式が離散型選択モデル、(2.12)式と(2.14)式がLISRELモデルを構成している。

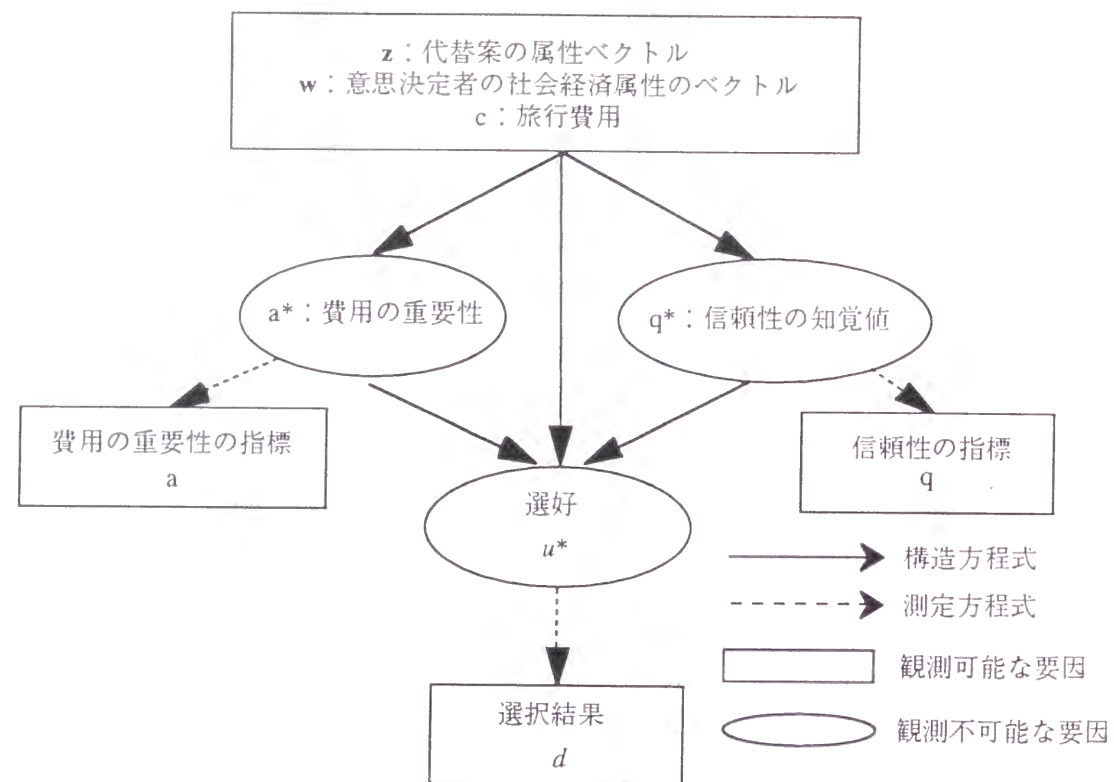


図 2-4 具体例の概略

ここで示したフレームワークをモデル化した具体的例を図 2-4 に示す。このパスダイヤグラムは、図 2-1 で示したものの 1 例である。ここでは代替案の数は 2 つで、効用関数は、信頼性という潜在変数と旅行費用及び代替案の他の属性で表されとし、費用に対する重みは意思決定者ごとに「費用の重要性」という潜在変数によって表されているとする。この例では、図 2-1 の「知覚値」が「信頼性」に、「態度」が「費用の重要性」にそれぞれ相当している。

c：旅行費用

z：代替案の属性のベクトル

q\*：信頼性の知覚値

q：アンケートなどで得られる信頼性の指標

w：意思決定者の社会経済的属性のベクトル

a\*：費用の重要性

a：アンケートなどで得られる費用の重要性の指標

u\*：2 つの代替案の効用差

d：実際の選択（1 は代替案 1，-1 は代替案 2）

(\* の付いているものは、潜在的であることを示す)

を示している、とする。

変数相互の関係は、下記の構造方程式で表される。

$$q^* = \beta_1' z + \zeta_1 \quad (2.15)$$

$$a^* = \beta_2' w + \zeta_2 \quad (2.16)$$

$$u^* = (\gamma_1 + \gamma_2 a^*)c + \gamma_3 q^* + \Gamma' z + v \quad (2.17)$$

指標と潜在変数の関係は、下記の測定方程式で表される。

$$q = \lambda_1 q^* + \varepsilon_1 \quad (2.18)$$

$$a = \lambda_2 a^* + \varepsilon_2 \quad (2.19)$$

ただし、 $\Gamma, \beta_1, \beta_2$  は、未知パラメータのベクトルで、 $\gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2$  は未知パラメータ、 $\zeta_1, \zeta_2, v, \varepsilon_1, \varepsilon_2$  はランダム項である。

(2.15)、(2.16) 式は、潜在変数と観測可能な変数の関係を表し、(2.18)、(2.19) 式は、潜在的要因の観測可能な指標を潜在的要因の関数として表している。これら 4 方程式は、LISREL モデルである。

ここでは、2-3-3 で示したフレームワークの推定方法について、同時推定と段階推定の 2 つの方法を提示する。またその 2 つのなかの改良の方向として、認知的不協和を考慮したモデルと、知覚値指標の離散性を考慮したモデルを提案する。

## 2-4 推定方法の検討と数理心理データの特性を考慮した定式化

### 2-4-1 段階推定

#### 1) 選択確率の誘導

2-3-3 で示したすべての変数が正規分布に従うと仮定すると以下のような誘導が行なわれる。 $Y, x^*, u^*$  の同時確率分布は、

$$\begin{bmatrix} Y \\ x^* \\ u^* \end{bmatrix} \sim \text{MVN}(\mathbf{M}_1, \mathbf{\Omega}_1) \quad (2.20)$$

ただし、

$$\mathbf{M}_1 = \begin{bmatrix} \Lambda \mathbf{B} \mathbf{s} \\ \mathbf{B} \mathbf{s} \\ \mathbf{a}' \mathbf{x} + \mathbf{c}' \mathbf{B} \mathbf{s} \end{bmatrix} \quad (2.21)$$

$$\mathbf{\Omega}_1 = \begin{bmatrix} \Lambda \Psi \Lambda' + \Theta & \Lambda \Psi & \Lambda \Psi \mathbf{c} \\ \Psi \Lambda' & \Psi & \Psi \mathbf{c} \\ \mathbf{c}' \Psi \Lambda' & \mathbf{c}' \Psi & 1 + \mathbf{c}' \Psi \mathbf{c} \end{bmatrix} \quad (2.22)$$

である。ここで、観測可能な変数  $\mathbf{Y}, \mathbf{x}, \mathbf{s}$  が与えられたときの  $\mathbf{x}^*, u^*$  の条件付き分布は、

$$\begin{bmatrix} \mathbf{x}^* \\ u^* \end{bmatrix} \sim \text{MVN}(\mathbf{M}_2, \mathbf{\Omega}_2) \quad (2.23)$$

ただし、

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{B} \mathbf{s} + \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} (\mathbf{Y} - \Lambda \mathbf{B} \mathbf{s}) \\ \mathbf{a}' \mathbf{x} + \mathbf{c}' \{ \mathbf{B} \mathbf{s} + \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} (\mathbf{Y} - \Lambda \mathbf{B} \mathbf{s}) \} \end{bmatrix} \quad (2.24)$$

および、

$$\boldsymbol{\omega} = \Psi - \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} \Lambda \Psi \quad (2.25)$$

と定義することによって、

$$\mathbf{\Omega}_2 = \begin{bmatrix} \boldsymbol{\omega} & \boldsymbol{\omega} \mathbf{c} \\ \mathbf{c}' \boldsymbol{\omega} & 1 + \mathbf{c}' \boldsymbol{\omega} \mathbf{c} \end{bmatrix} \quad (2.26)$$

となる。

このとき、離散型選択モデルの選択確率は次式で与えられる。

$$\Pr(d \mid \mathbf{x}, \mathbf{Y}, \mathbf{s}) = \Phi \left( d \frac{\mathbf{a}' \mathbf{x} + \mathbf{c}' \{ \mathbf{B} \mathbf{s} + \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} (\mathbf{Y} - \Lambda \mathbf{B} \mathbf{s}) \}}{\sqrt{1 + \mathbf{c}' \boldsymbol{\omega} \mathbf{c}}} \right) \quad (2.27)$$

ただし、変数の定義は、全て2章で示したものと同一とする。

## 2) 推定法

線形構造方程式モデルと、離散型選択モデルを段階的に推定する方法は、まず(2.12)式と(2.14)式で構成される線形構造方程式モデルを LISREL などのプログラムで推定し、そのパラメータ推定値を用いて潜在変数  $\mathbf{x}^*$  の推計値 (fitted value) を計算し、その値を(2.11)式の効用関数に代入したうえで(2.11)式、(2.13)式で表される離散型選択モデルを推定するという

ものである。つまり、最終的には(2.27)式で表される選択確率によってプロビット・モデルを推定することになる。この段階推定により、同時推定量に比べ有効性は落ちるが一致性のある推定量を得ることができる (Amemiya (1985))。具体的には次のような2つのステップを行う。

### Step 1

LISREL などのプログラムにより線形構造方程式モデルのパラメータを推定し、次式により潜在変数  $\mathbf{x}^*$  およびその分散の推計値を計算する。

$$\widehat{\mathbf{x}}^* = \widehat{\mathbf{B}} \mathbf{s} + \widehat{\Psi} \widehat{\Lambda}' (\widehat{\Lambda} \widehat{\Psi} \widehat{\Lambda}' + \widehat{\Theta})^{-1} (\mathbf{y} - \widehat{\Lambda} \widehat{\mathbf{B}} \mathbf{s}) \quad (2.28)$$

$$\widehat{\boldsymbol{\omega}} = \widehat{\Psi} - \widehat{\Psi} \widehat{\Lambda}' (\widehat{\Lambda} \widehat{\Psi} \widehat{\Lambda}' + \widehat{\Theta})^{-1} \widehat{\Lambda} \widehat{\Psi} \quad (2.29)$$

### Step 2

選択モデルの未知パラメータ  $\mathbf{a}, \mathbf{c}$  を選択確率が次式で表される修正プロビット・モデルによって推定する。

$$\Pr(d \mid \mathbf{x}, \mathbf{Y}, \mathbf{s}) = \Phi \left( d \frac{\mathbf{a}' \mathbf{x} + \mathbf{c}' \widehat{\mathbf{x}}^*}{\sqrt{1 + \mathbf{c}' \widehat{\boldsymbol{\omega}} \mathbf{c}}} \right) \quad (2.30)$$

## 2-4-2 認知的不協和 (cognitive dissonance) を考慮した段階推定

### 1) 認知的不協和

2-2-3 で述べたように、本研究では認知的不協和の存在しうるデータを用いるため、認知的不協和を考慮したモデルの定式化を行う。ここで再び認知的不協和を簡単に説明すると、認知的不協和とは自己の選択した代替案に対して、その要因を好意的に解釈しようとするもので、自己の行動の正当化の一種であると考えられる。本研究は、RP (Revealed Preference) データ、つまり被験者の実際の選択結果が明らかな行動に対して取られたデータであるため、アンケートの中で代替案の主観的評価を尋ねた項目に、自分の選択した結果とアンケートの回答が矛盾しないよう選択したモードを過大に評価する可能性があるため、主観的評価値にバイアスが生じて LISREL モデルの推定結果の信頼性が低下するおそれがあるということである。

そのようなバイアスを除去するためのモデルを以下のように定式化する。



### 構造方程式

$$u^* = \mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{x}^* + v \quad (2.31)$$

$$\mathbf{x}^* = \mathbf{B}\mathbf{s} + \zeta \quad (2.32)$$

### 測定方程式

$$d = \begin{cases} 1 & \text{if } u^* \geq 0 \\ -1 & \text{if } u^* < 0 \end{cases} \quad (2.33)$$

$$\mathbf{Y} = \mathbf{\Lambda}\mathbf{x}^* + \mathbf{\Gamma}d + \varepsilon \quad (2.34)$$

ただし、

$\mathbf{\Gamma}$  = 未知パラメータの行列

(2.34) 式に示すように、知覚値の測定方程式に実際の選択結果を反映させることで、先に述べた選択結果が知覚値指標に影響を与えているという効果を測定することことができる。

### 2) 選択確率の誘導

$d$  を  $\mathbf{B}$  に係数をもたず  $\mathbf{\Lambda}$  にのみ係数をもつ完全に計量された潜在変数とすることで、LISREL モデルの一般形にあてはめることができる。つまり  $\mathbf{\Gamma}$  は、 $\mathbf{\Lambda}$  の一部として考えることができ、選択確率は通常の段階推定と同様に導かれる。

### 3) 推定方法

上記で述べたモデルを推定する方法は、通常の段階推定と同様に線形構造方程式モデルを LINES などのプログラムで推定し、そのパラメータ推定値を用いて定性的変数  $\mathbf{x}^*$  の推計値 (fitted value) を計算するのであるが、認知的不協和を取り除くために導入したパラメータ  $\mathbf{\Gamma}$  は、知覚値指標のバイアスを表していると考えられ、それを除外したものが真のパラメータを示していると考えられる。よって  $\mathbf{x}^*$  の推計値 (fitted value) を計算するときは、パラメータ  $\mathbf{\Gamma}$  を除いたパラメータ行列で計算し、その値を (2.11) 式の効用関数に代入したうえで (2.11) 式、(2.13) 式で表される離散型選択モデルを推定するというものである。

## 2-4-3 同時推定

### 1) 選択確率の誘導

段階推定と同様に、すべての変数が正規分布に従うと仮定すると以下のような誘導が行われる。 $\mathbf{Y}$ 、 $\mathbf{x}^*$ 、 $u^*$  の同時確率分布は、

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{x}^* \\ u^* \end{bmatrix} \sim \text{MVN}(\mathbf{M}_1, \mathbf{\Omega}_1) \quad (2.35)$$

ただし、

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{\Lambda}\mathbf{B}\mathbf{s} \\ \mathbf{B}\mathbf{s} \\ \mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{B}\mathbf{s} \end{bmatrix} \quad (2.36)$$

$$\mathbf{\Omega}_1 = \begin{bmatrix} \mathbf{\Lambda}\Psi\mathbf{\Lambda}' + \theta & \mathbf{\Lambda}\Psi & \mathbf{\Lambda}\Psi\mathbf{c} \\ \Psi\mathbf{\Lambda}' & \Psi & \Psi\mathbf{c} \\ \mathbf{c}'\Psi\mathbf{\Lambda}' & \mathbf{c}'\Psi & 1 + \mathbf{c}'\Psi\mathbf{c} \end{bmatrix} \quad (2.37)$$

である。

ここで、 $\mathbf{x}^*$  が与えられたときの  $\mathbf{Y}$ 、 $u^*$  の条件付き分布は、

$$\begin{bmatrix} \mathbf{Y} \\ u^* \end{bmatrix} \sim \text{MVN}(\mathbf{M}_2, \mathbf{\Omega}_2) \quad (2.38)$$

ただし、

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{\Lambda}\mathbf{x}^* \\ \mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{x}^* \end{bmatrix} \quad (2.39)$$

$$\mathbf{\Omega}_2 = \begin{bmatrix} \Theta & 0 \\ 0 & 1 \end{bmatrix} \quad (2.40)$$

となる。このとき  $\mathbf{Y}$ 、 $d$  の同時生起確率は  $\mathbf{x}^*$  の分布形を与えることによって、次式で与えられる。

$$\Pr(d, \mathbf{Y}) = \iint_{\mathbf{x}^*} \Phi\left(d(\mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{x}^*)\right) \prod_{i=1}^6 \phi\left(\frac{Y_i - (\lambda_{i1}x_1^* + \lambda_{i2}x_2^*)}{\theta_i}\right) \cdot \prod_{j=1}^2 \phi\left(\frac{x_j^* - \mathbf{B}\mathbf{s}}{\psi_j}\right) d\mathbf{x}^* \quad (2.41)$$

## 2) 推定法

線形構造方程式モデルと離散型選択モデルを同時に推定する方法は、 $\mathbf{x}^*$  が平均値  $\mathbf{B}\mathbf{s}$ 、分散  $\Psi$  の正規分布に従うとしたときの  $\mathbf{Y}$ 、 $d$  の条件付き分布が (2.38) 式で与えられるので  $\mathbf{Y}$ 、 $d$  の条件付き同時生起確率に  $\mathbf{x}^*$  の確率密度関数をかけ、さらに  $\mathbf{x}^*$  について積分を行なうことにより  $\mathbf{Y}$ 、 $d$  の出現確率が得られる。ただしパラメータ同定のため  $\Psi$  は、 $\mathbf{I}$  ( $\mathbf{I}$  は単位行列) であるとする。この同時推定により得られる最尤推定量は、段階推定より漸近的有効性が高く一致性をもつ。

### 2-4-4 知覚値指標の離散性を考慮した同時推定

ここまでの定式化では、知覚値指標  $\mathbf{Y}$  を 1「悪い」、2「やや悪い」、3「普通」、4「やや良い」、5「良い」の選択肢から選ばれたアンケート結果を、そのまま連続変数として用いている。しかし厳密には知覚値自体は連続変数であるが、得られたアンケート結果は離散的であって、その閾値の間隔は一定ではないと考えられる。つまり、「普通」という回答が生じる閾値の間隔と、「やや良い」という閾値の間隔は異なると考えられる。この問題を解決するため、知覚値指標の離散性を考慮したモデルの推定方法を述べる。

#### 1) 選択確率の誘導

真の知覚値を  $\mathbf{Y}^*$  とし、すべての変数が正規分布に従うと仮定すると以下のような誘導が行われる。 $\mathbf{Y}^*$ 、 $\mathbf{x}^*$ 、 $u^*$  の同時確率分布は、

$$\begin{bmatrix} \mathbf{Y}^* \\ \mathbf{x}^* \\ u^* \end{bmatrix} \sim \text{MVN}(\mathbf{M}_1, \mathbf{\Omega}_1) \quad (2.42)$$

ただし、

$$\mathbf{M}_1 = \begin{bmatrix} \Lambda \mathbf{B} \mathbf{s} \\ \mathbf{B} \mathbf{s} \\ \mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{B} \mathbf{s} \end{bmatrix} \quad (2.43)$$

$$\mathbf{\Omega}_1 = \begin{bmatrix} \Lambda \Psi \Lambda' + \Theta & \Lambda \Psi & \Lambda \Psi \mathbf{c} \\ \Psi \Lambda' & \Psi & \Psi \mathbf{c} \\ \mathbf{c}' \Psi \Lambda' & \mathbf{c}' \Psi & 1 + \mathbf{c}' \Psi \mathbf{c} \end{bmatrix} \quad (2.44)$$

ここで、 $\mathbf{Y}^*$  と  $\mathbf{Y}$  の関係について

$$Y_i = j \text{ if } \theta_{j-1} < Y_i^* \leq \theta_j \text{ for } j=0, \dots, 6 \quad \theta_0 = -\infty, \theta_6 = +\infty \quad (2.45)$$

$$\mathbf{A} = \{\mathbf{Y}^* | \text{consistent with observed } \mathbf{Y}\} \quad (2.46)$$

として、 $\mathbf{x}^*$  が与えられたときの  $\mathbf{Y}^*$  の条件付き分布は、

$$\mathbf{Y}^* | \mathbf{x}^* = N(\Lambda \mathbf{x}^*, \Theta) \quad (2.47)$$

で与えられる。これから、 $\mathbf{x}^*$  が与えられたときの  $\mathbf{A}$  の尤度は、

$$\prod_{i=1}^6 \left[ \Phi\left(\frac{\theta_{j_i} - (\lambda_{i1}x_1^* + \lambda_{i2}x_2^*)}{\theta_i}\right) - \Phi\left(\frac{\theta_{j_i-1} - (\lambda_{i1}x_1^* + \lambda_{i2}x_2^*)}{\theta_i}\right) \right] \quad (2.48)$$

となり、このとき  $Y_i = j_i$  ならば、 $\mathbf{s}$  が与えられたときの  $\mathbf{Y}$  の尤度は

$$\iint_{\mathbf{x}^*} \prod_{i=1}^6 \left[ \Phi\left(\frac{\theta_{j_i} - (\lambda_{i1}x_1^* + \lambda_{i2}x_2^*)}{\theta_i}\right) - \Phi\left(\frac{\theta_{j_i-1} - (\lambda_{i1}x_1^* + \lambda_{i2}x_2^*)}{\theta_i}\right) \right] \cdot \prod_{j=1}^2 \phi\left(\frac{x_j^* - \mathbf{B}\mathbf{s}}{\psi_j}\right) d\mathbf{x}^* \quad (2.49)$$

となる。

また (2.42) 式から、 $\mathbf{x}^*$  が与えられたときの  $\mathbf{Y}^*$  と  $u^*$  の条件付き分布が、

$$\begin{bmatrix} \mathbf{Y}^* \\ u^* \end{bmatrix} \sim \text{MVN}(\mathbf{M}_2, \mathbf{\Omega}_2) \quad (2.50)$$

ただし、

$$\mathbf{M}_2 = \begin{bmatrix} \Lambda \mathbf{x}^* \\ \mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{x}^* \end{bmatrix} \quad (2.51)$$

$$\Omega_2 = \begin{bmatrix} \Theta & 0 \\ 0 & 1 \end{bmatrix} \quad (2.52)$$

と求められる。

ここで選択モード  $d$  と  $\theta_j < Y_i \leq \theta_{j+1}$  となるような  $d_j$  が観測されたときの  $\mathbf{Y}$ 、 $d$  の同時生起確率は、 $\mathbf{x}^*$  の分布形を与えることによって、次式で与えられる。

$$Pr(d, \mathbf{Y}) = \iint_{\mathbf{x}^*} \Phi\left(d(\mathbf{a}'\mathbf{x} + \mathbf{c}'\mathbf{x}^*)\right) \prod_{i=1}^6 \left[ \Phi\left(\frac{\theta_{\delta_i+1} - (\lambda_{i1}\mathbf{x}_1^* + \lambda_{i2}\mathbf{x}_2^*)}{\theta_i}\right) - \Phi\left(\frac{\theta_{\delta_i} - (\lambda_{i1}\mathbf{x}_1^* + \lambda_{i2}\mathbf{x}_2^*)}{\theta_i}\right) \right] \cdot \prod_{j=1}^2 \phi\left(\frac{\mathbf{x}_j^* - \mathbf{B}\mathbf{s}}{\Psi_j}\right) d\mathbf{x}^* \quad (2.53)$$

## 2) 推定方法

知覚値指標の離散性を考慮したモデルの推定方法は、同時推定と同様に、 $\mathbf{x}^*$  が平均値  $\mathbf{B}\mathbf{s}$ 、分散  $\Psi$  の正規分布に従うとしたときの  $\mathbf{Y}^*$ 、 $u^*$  の条件付き分布が (2.50) 式で与えられる。また同じ条件下での観測された、知覚値指標  $\mathbf{Y}$  と一致する真の知覚値  $\mathbf{Y}^*$  の関係が (2.49) 式で与えられているから、 $\mathbf{Y}$ 、 $d$  の条件付き同時生起確率に  $\mathbf{x}^*$  の確率密度関数をかけ、さらに  $\mathbf{x}^*$  について積分を行なうことにより  $\mathbf{Y}$ 、 $d$  の出現確率が得られる。なおパラメータ同定のため  $\Psi$  は、 $\mathbf{I}$  ( $\mathbf{I}$  は単位行列) であるとする。

この考え方にに基づき最尤推定法を用いた推定量は、漸近的有効性を持ち、一致性がある。

## 2-5 事例研究

### 2-5-1 データの概略

事例研究で用いるデータは、1987 年にオランダで行われた都市間旅行において列車と自動車の機関選択に関するアンケート調査によっている。このアンケート調査は西ドイツとの国境に近いナイメゲン (Nijmegen) という都市の住民に対するもので、この都市から列車または車どちらでも約 2 時間の距離にある、アムステルダム、ロッテルダム、デン・ハーグなどを中核都市とするランドシュタット (Randstad) と呼ばれる大都市圏への旅行を対象としている。調査対象は、過去 3 カ月間にランドシュタットに行ったことがあり、列車の定期券を持っておらず、車も列車も支障なく使用が可能な人々で、標本抽出は、電話帳からランダムサンプリングをして、

家庭訪問調査が可能であるかどうかを尋ねた上で、上記の条件を満たしている 235 人が選ばれた。実際には、その中からデータに不備のない 219 人のデータを使用した。

質問項目は、実際に行なったランドシュタットへの旅行に対して旅行費用・旅行時間などのトリップ属性、個人の社会経済属性を含み、トリップ属性に関しては、選択したモードおよび選択しなかったモードについてそれぞれ認識している値の回答を得ている。また、トリップ属性に関する主観的評価値として次の 6 項目を選択モード、非選択モードに対して回答者に尋ねている。( ) 内は後の定式化のときの変数名を示す。

- 1) 旅行中の安楽度 (*Relax*)
- 2) 到着時刻の信頼性 (*Relia*)
- 3) 出発時刻の柔軟性 (*Flex*)
- 4) 大きな荷物があるときや、子供を連れているときの利用しやすさ (*Ease*)
- 5) 旅行中の安全性 (*Safe*)
- 6) 全体として (*Overall*)

その回答は 1) ~ 5) が 1) 非常に悪い、2) 悪い、3) 普通、4) 良い、5) 非常に良いの 5 段階で評価し、6) は 10 段階に評点を付けさせている。実際にこのアンケートは、RP データだけでなく、仮想の列車に対する SP データが、2 種類取られている。そのデータに関しては、山田 (1991) に詳しい。

### 2-5-2 モデルの特定化

本データの分析では、主観的評価指標の数と内容から、旅行中の快適性及び交通機関の利便性という 2 つの潜在的要因を考慮にいて定式化することにした。なお、式中では主観的評価値を含むトリップ属性変数は鉄道の値から自動車の値を引いたものになっている。以下、特定化に使用した変数の定義を述べる。

|                |   |                                   |
|----------------|---|-----------------------------------|
| <i>aged</i>    | = | 1 : 40 歳以上の時 ; 0 : そうでないとき        |
| <i>ltime</i>   | = | 幹線旅行時間 (乗り換え時間を含む) (単位 : 時間)      |
| <i>first</i>   | = | 1 : 鉄道で 1 等車を利用するとき ; 0 : そうでないとき |
| <i>trmtime</i> | = | 端末旅行時間 (単位 : 時間)                  |
| <i>xfern</i>   | = | 鉄道を利用したときの乗り換え回数                  |

freepark = 1：目的地で無料駐車ができるとき；0：そうでないとき  
 costpp = 一人当たり旅行費用（Guilder）  
 business = 1：ビジネスに関する旅行のとき；0：そうでないとき  
 female = 1：女性；0：男性

#### 1）段階推定

LISREL モデルの構造方程式を次のように特定化した。

旅行中の快適性

$$x_1^* = \beta_1 \text{aged} + \beta_2 \text{lh time} + \beta_3 \text{first} + \beta_4 \text{aged} \times \text{lh time} + \zeta_1 \quad (2.54)$$

交通機関の利便性

$$x_2^* = \beta_5 \text{aged} + \beta_6 \text{tr mtime} + \beta_7 \text{xfern} + \beta_8 \text{freepark} + \zeta_2 \quad (2.55)$$

これら2つの潜在変数と主観的評価値を関係付ける測定方程式を次のように定めた。

$$\begin{bmatrix} y_1 (\text{relax}) \\ y_2 (\text{reliability}) \\ y_3 (\text{flex}) \\ y_4 (\text{ease}) \\ y_5 (\text{safe}) \\ y_6 (\text{overall}) \end{bmatrix} = \begin{bmatrix} 1 & \lambda_4 \\ \lambda_1 & 1 \\ 0 & \lambda_5 \\ 0 & \lambda_6 \\ \lambda_2 & \lambda_7 \\ \lambda_3 & \lambda_8 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (2.56)$$

上式の未知パラメーター行列では、パラメーター同定のために各列で1つのパラメーターが1に正規化されている。

次に、離散型選択モデルの構造方程式（効用関数）を次のように特定化した。

$$u^* = a_0 + a_1 \text{costpp} + a_2 \text{lh time} + a_3 \text{tr mtime} + a_4 \text{xfern} + a_5 \text{business} + a_6 \text{female} + c_1 x_1^* + c_2 x_2^* + v \quad (2.57)$$

#### 2）認知的不協和を考慮した段階推定

同時推定と基本的に同一であるが、アンケート結果の中で特に認知的不協和の影響を受けやすいとおもわれる  $y_6$  (overall) にバイアス除去のための選択結果ダミーのパラメーターが入る。

$$\begin{bmatrix} y_1 (\text{relax}) \\ y_2 (\text{reliability}) \\ y_3 (\text{flex}) \\ y_4 (\text{ease}) \\ y_5 (\text{safe}) \\ y_6 (\text{overall}) \end{bmatrix} = \begin{bmatrix} 1 & \lambda_4 & 0 \\ \lambda_1 & 1 & 0 \\ 0 & \lambda_5 & 0 \\ 0 & \lambda_6 & 0 \\ \lambda_2 & \lambda_7 & 0 \\ \lambda_3 & \lambda_8 & \lambda_9 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ d \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (2.58)$$

#### 3）同時推定

旅行中の快適性

$$x_1^* = \beta_1 \text{aged} + \beta_2 \text{lh time} + \beta_3 \text{first} + \beta_4 \text{aged} \times \text{lh time} + \zeta_1 \quad (2.59)$$

交通機関の利便性

$$x_2^* = \beta_5 \text{aged} + \beta_6 \text{tr mtime} + \beta_7 \text{xfern} + \beta_8 \text{freepark} + \zeta_2 \quad (2.60)$$

また測定方程式を次のように定める。

$$\begin{bmatrix} y_1 (\text{relax}) \\ y_2 (\text{reliability}) \\ y_3 (\text{flex}) \\ y_4 (\text{ease}) \\ y_5 (\text{safe}) \\ y_6 (\text{overall}) \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_5 \\ \lambda_2 & \lambda_6 \\ 0 & \lambda_7 \\ 0 & \lambda_8 \\ \lambda_3 & \lambda_9 \\ \lambda_4 & \lambda_{10} \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (2.61)$$

ここでは、モデル同定のために各列で1つパラメーターを1に固定せず、構造方程式のランダム項の分散を1に固定した。

次に離散型選択モデルの構造方程式（効用関数）を次のように特定化した。

$$u^* = a_0 + a_1 \text{costpp} + a_2 \text{lh time} + a_3 \text{tr mtime} + a_4 \text{xfern} + a_5 \text{business} + a_6 \text{female} + c_1 x_1^* + c_2 x_2^* + v \quad (2.62)$$

#### 4）知覚値指標の離散性を考慮した同時推定

同時推定モデルと同様の特定化を行なった。



2-5-3 各推定結果の考察

1) 段階推定

LISREL モデルを最尤推定法で推定した結果を表 2-2 に示す。またこの結果から潜在変数が (2.28) 式によって計算される。この値を用いて選択モデルを推定した結果を表 2-3 に示す。

表 2-2 の結果をみると、その推定値は t 値が小さいものも多いがいずれも予想通りの符号を持ち、2 つの潜在変数は心因的快適性 (Comfort\*) と利便性 (Convenience\*) を表していると考えられる。また、潜在変数の有効性を検証するために、潜在変数を用いないモデルの推定結果を同じ表中に示した。表 2-3 の結果をみると、LISREL モデルによって求められた潜在変数は、いずれも有意な正の係数を持ち、2 つのモデルを比較すると、McFadden の提唱する適合度を示す指標である  $\overline{\rho^2}$  (Ben-Akiva and Lerman (1985)) の値より、潜在変数を用いたモデルは潜在変数を用いないモデルよりも適合度が大きく上がっていることがわかる。

推定結果を詳しくみると、潜在変数を用いることで幹線旅行時間の係数が、正から負へと変化しているが、これは、幹線旅行時間が、潜在変数 Comfort\* に用いられているため、重共線性の問題が起きたと考えられる。また、端末旅行時間の係数の絶対値が小さくなりかつ t 値も小さくなっていることも、同じように端末旅行時間は潜在変数 Convenience\* のなかに用いられていることによると考えられる。また幹線旅行時間がほぼ 0 になっているということは、幹線旅行時間は、選択行動において Comfort\* を通じてのみ影響を与えているためとも考えられる。

表 2-2 段階推定による LISREL モデルの推定結果

|              |           |           |                 |
|--------------|-----------|-----------|-----------------|
| $\hat{B}' =$ | $(x_1^*)$ | $(x_2^*)$ |                 |
|              | -0.232    | 0.406     | (aged)          |
|              | (-1.4)    | (3.3)     |                 |
|              | -0.292    | 0         | (lhtime)        |
|              | (-1.3)    |           |                 |
|              | 0         | -0.522    | (trmtime)       |
|              |           | (-2.1)    |                 |
|              | 0.286     | 0         | (first)         |
|              | (1.0)     |           |                 |
| $\hat{A} =$  | 0         | -0.0471   | (xfern)         |
|              |           | (-0.6)    |                 |
|              | 0         | 0.164     | (freepark)      |
|              |           | (1.6)     |                 |
|              | -0.0405   | 0         | (aged × lhtime) |
|              | (-0.1)    |           |                 |

( ) 内は t 値

表 2-3 段階推定による選択モデルの推定結果

| 変数名                 | 潜在変数を用いたモデル    | 潜在変数を用いないモデル   |
|---------------------|----------------|----------------|
| 鉄道定数                | 0.322 (1.0)    | 0.538 (2.0)    |
| 費用                  | -0.0338 (-4.1) | -0.0268 (-4.2) |
| 幹線旅行時間              | 0.0751 (0.2)   | -0.405 (-1.6)  |
| 端末旅行時間              | -1.18 (-2.6)   | -1.57 (-4.2)   |
| 乗換回数                | -0.316 (-1.7)  | -0.195 (-1.3)  |
| ビジネスダミー             | 1.33 (3.6)     | 0.942 (3.4)    |
| 女性ダミー               | 0.652 (2.6)    | 0.466 (2.3)    |
| 快適性*                | 0.882 (2.7)    |                |
| 利便性*                | 1.39 (4.1)     |                |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.352          | 0.242          |

( ) 内は t 値

また定数項が有意な係数であったものが、潜在変数を用いることでその有意性を失っているのは、本来必要であった潜在変数 Comfort\* と Convenience\* が抜けていたためにそれを補正するためのバイアス (Omitted Variable Bias) が生じていたためと考えられる。その他の説明変数の係数の絶対値が大きくなっていることは、ランダム効用の分散が小さくなっていることを示していると考えられる。

2) 認知的不協和を考慮した段階推定

LISREL モデルの推定結果を表 2-4 に示す。同時推定と同様の結果を示しているが、各未知パラメータの推定値の中で、構造方程式中の aged × lhtime の係数が正になっているのは、その有意性の低さからほぼ 0 であるものが、たまたま正の値に推定されたためと思われる。またその他の係数は予想通りの符号をもっているが、いずれもその t 値は低い。また、測定方程式中の選択結果ダミーのパラメータの推定値が有意な正の値になっている。つまり、被験者は知覚値指標アンケートの (overall) の項目が、モデルで仮定した構造とは別に選択した交通機関と有意に正の相関を持っていることを示しており、選択した交通機関を意識的に良く評価するとい

表 2－4 認知的不協和を考慮した段階推定の LISREL モデルの推定結果

$$\hat{\mathbf{B}}' = \begin{bmatrix} (x_1^*) & (x_2^*) \\ -0.173 & 0.378 \\ (-0.5) & (1.8) & (aged) \\ -0.370 & 0 \\ (-1.5) & & (lhtime) \\ 0 & -0.248 \\ & (-0.8) & (trmtime) \\ 0.147 & 0 \\ (0.5) & & (first) \\ 0 & -0.0017 \\ & (-0.01) & (xfern) \\ 0 & 0.130 \\ & (1.2) & (freepark) \\ 0.0760 & 0 \\ (0.2) & & (aged \times lhtime) \end{bmatrix} \quad \hat{\mathbf{\Lambda}} = \begin{bmatrix} (x_1^*) & (x_2^*) & (d) \\ 1 & 0.224 & 0 \\ & (0.4) & (relax) \\ 1.00 & 1 & 0 \\ (1.2) & & (relia) \\ 0 & 1.98 & 0 \\ & (2.0) & (flex) \\ 0 & 1.12 & 0 \\ & (2.6) & (ease) \\ 0.593 & 0.295 & 0 \\ (2.0) & (0.8) & (safe) \\ 1.51 & 2.31 & 1 \\ (1.6) & (3.1) & (overall) \end{bmatrix}$$

( ) 内は t 値

う認知的不協和の影響による可能性が高いと思われる。

その選択結果ダミーのパラメータを取り除いた LISREL モデルの推定結果から計算された潜在変数を用いた選択モデルの推定結果を表 2－5 に示す。ここでも、1) の段階推定と同様の結果が生じている。

表 2－5 認知的不協和を考慮した段階推定による選択モデルの推定結果

| 変数名                 | 潜在変数を用いたモデル    | 潜在変数を用いないモデル   |
|---------------------|----------------|----------------|
| 鉄道定数                | 0.566 (1.2)    | 0.538 (2.0)    |
| 費用                  | -0.0495 (-3.4) | -0.0268 (-4.2) |
| 幹線旅行時間              | 0.178 (0.4)    | -0.405 (-1.6)  |
| 端末旅行時間              | -1.91 (-2.7)   | -1.57 (-4.2)   |
| 乗換回数                | -0.477 (-1.7)  | -0.195 (-1.3)  |
| ビジネスダミー             | 1.92 (3.3)     | 0.942 (3.4)    |
| 女性ダミー               | 0.958 (2.4)    | 0.466 (2.3)    |
| 快適性*                | 1.52 (2.3)     |                |
| 利便性*                | 2.08 (3.2)     |                |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.347          | 0.242          |

( ) 内は t 値

### 3) 同時推定

表 2－6 に LISREL モデル部分の推定結果を表 2－7 に選択モデル部分の推定結果を示す。各パラメータは予想どうりの符号を持ち、2つの潜在変数は、心因的快適性 (Comfort<sup>\*</sup>) と利便

表 2－6 同時推定による LISREL モデルの推定結果

$$\hat{\mathbf{B}}' = \begin{bmatrix} (x_1^*) & (x_2^*) \\ -0.427 & 0.378 \\ (-2.4) & (2.4) & (aged) \\ -0.323 & 0 \\ (-1.7) & & (lhtime) \\ 0 & -1.98 \\ & (-9.0) & (trmtime) \\ 0.281 & 0 \\ (0.9) & & (first) \\ 0 & -0.396 \\ & (-3.7) & (xfern) \\ 0 & 0.482 \\ & (3.5) & (freepark) \\ -0.339 & 0 \\ (-1.3) & & (aged \times lhtime) \end{bmatrix} \quad \hat{\mathbf{\Lambda}} = \begin{bmatrix} (x_1^*) & (x_2^*) \\ 0.433 & 0.280 \\ (7.6) & (3.2) & (relax) \\ 0.527 & 0.661 \\ (12.5) & (10.2) & (relia) \\ 0 & 0.815 \\ & (14.7) & (flex) \\ 0 & 0.794 \\ & (14.2) & (ease) \\ 0.462 & 0.311 \\ (11.6) & (5.2) & (safe) \\ 0.784 & 1.76 \\ (8.5) & (14.1) & (overall) \end{bmatrix}$$

( ) 内は t 値

表 2－7 同時推定による選択モデルの推定結果

| 変数名                 | 潜在変数を用いたモデル    | 潜在変数を用いないモデル   |
|---------------------|----------------|----------------|
| 鉄道定数                | -1.81 (-0.9)   | 0.538 (2.0)    |
| 費用                  | -0.0379 (-4.3) | -0.0268 (-4.2) |
| 幹線旅行時間              | 0.379 (0.9)    | -0.405 (-1.6)  |
| 端末旅行時間              | -0.818 (-2.3)  | -1.57 (-4.2)   |
| 乗換回数                | -0.230 (-1.2)  | -0.195 (-1.3)  |
| ビジネスダミー             | 1.28 (3.3)     | 0.942 (3.4)    |
| 女性ダミー               | 0.700 (2.9)    | 0.466 (2.3)    |
| 快適性*                | 1.29 (-1.8)    |                |
| 利便性*                | 1.10 (4.7)     |                |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.201※         | 0.242          |

※LISREL モデルとあわせた全体の尤度比 ( ) 内は t 値

性（Convenience<sup>\*</sup>）を表していると考えられる。特に LISREL モデル部分のパラメターの t 値が大きいものが多い。選択モデル部分の結果をみると、定数項の値の絶対値が大きくなっていることに気づく。これは、その t 値が低いことから考えて、推定値自体はたまたま大きくなっているが、実際には 0 である確率が大きいことを示している。また Comfort<sup>\*</sup> の t 値があまり大きくないことと、幹線旅行時間の係数が正でかつその t 値が低いことは、Comfort<sup>\*</sup> に幹線旅行時間が含まれているため、その重共線性の影響と考えられる。

4) 知覚値指標の離散性を考慮した同時推定

LISREL モデル部分の推定結果を表 2-8 に、選択モデル部分の推定結果を表 2-9 にそれぞれ示す。ここでは、いままでと違い予想に反する符号を持つものが多い。これは、それまで連続変数として扱ってきたものを離散的なものとして扱うのであるから推定結果に大きな相違がみられたと考えられる。その LISREL モデル中の符号から考えて符号に多少の問題はあるが、 $x_1^*$  と  $x_2^*$  はそれぞれ快適性と（単純な）利便性を表していると考えられる。また閾値は潜在変数の期待値を 0 としてその偏差として示されており、そこに注意して見るとその間隔はほぼ一定であるが負の方向にずれている。これは、「普通」から「やや良い」はわずかな違いで変化するが、「普通」から「やや悪い」には簡単に変化しないということである。これらのことから、

表 2-8 知覚値指標の離散性を考慮した同時推定による LISREL モデルの推定結果

$$\hat{B}' = \begin{bmatrix} (x_1^*) & (x_2^*) \\ -0.372 & 2.41 \\ (-8.6) & (11.4) & (aged) \\ -0.512 & 0 \\ (-3.1) & & (lhtime) \\ 0 & -2.93 \\ & (-8.2) & (trmtime) \\ -0.279 & 0 \\ (-0.8) & & (first) \\ 0 & -0.263 \\ & (-2.0) & (xfern) \\ 0 & 0.318 \\ & (2.5) & (freepark) \\ 0.661 & 0 \\ (2.1) & & (aged \times lhtime) \end{bmatrix}$$
$$\hat{A} = \begin{bmatrix} (x_1^*) & (x_2^*) \\ -1.48 & -0.753 \\ (-5.6) & (-4.7) & (relax) \\ -0.765 & -0.688 \\ (-5.9) & (-6.4) & (relia) \\ 0 & 0.206 \\ & (2.2) & (flex) \\ 0 & 0.740 \\ & (5.3) & (ease) \\ 0.0695 & -0.315 \\ (0.8) & (-3.4) & (safe) \\ 0.482 & 1.16 \\ (5.0) & (9.3) & (overall) \end{bmatrix}$$
$$threshold' = \begin{bmatrix} -2.24 & -1.09 & 0.227 & 1.75 \\ (-21.5) & (-12.0) & (2.3) & (14.8) \end{bmatrix}$$

( ) 内は t 値

表 2-9 知覚値指標の離散性を考慮した同時推定による選択モデルの推定結果

| 変数名                 | 潜在変数を用いたモデル    | 潜在変数を用いないモデル   |
|---------------------|----------------|----------------|
| 鉄道定数                | -0.376 (-0.8)  | 0.538 (2.0)    |
| 費用                  | -0.0408 (-4.5) | -0.0268 (-4.2) |
| 幹線旅行時間              | -0.235 (-1.6)  | -0.405 (-1.6)  |
| 端末旅行時間              | 0.216 (0.3)    | -1.57 (-4.2)   |
| 乗換回数                | -0.146 (-0.8)  | -0.195 (-1.3)  |
| ビジネスダミー             | 1.34 (8.4)     | 0.942 (3.4)    |
| 女性ダミー               | 0.701 (4.5)    | 0.466 (2.3)    |
| 快適性*                | 0.739 (4.0)    |                |
| 利便性*                | 1.15 (8.5)     |                |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.071※         | 0.242          |

※LISREL モデルとあわせた全体の尤度比 ( ) 内は t 値

知覚値の指標は、ほぼ連続変数として扱うことが可能であるが、その誤差項の平均値は 0 ではなく、少し負の方向にあるということを考慮しなければならない。また LISREL モデルの中の推定値の t 値は、いずれも大きな値であって、推定結果の有意性は高い。選択モデル部分をみると、端末旅行時間の係数が正になっている。これは Convenience<sup>\*</sup> との重共線性のためと考えられる。また、乗り換え回数や、幹線旅行時間の t 値が低いのも同じ理由のためと考えられる。

2-5-4 推定結果の比較

以上事例研究として、4 つのモデルの推定結果を示した。本研究の最終的な狙いは潜在変数を用いた交通機関選択分析の実用化であり、提案したモデルの有効性を検討する意味からその比較を行なう。以下簡単のため段階推定モデルをモデル 1、認知的不協和を考慮した段階推定モデルをモデル 2、同時推定モデルをモデル 3、知覚値指標の離散性を考慮したモデルをモデル 4、と呼ぶ。



まず、段階的に推定するという同一のカテゴリーに所属する、モデル1とモデル2について比較する。LISRELモデルの結果を比較するためにLRS（Likelihood Ratio Statistics）の値を比べると

LRS（Model1）=66.7                   （39）  
LRS（Model2）=106.6               （39）（）内は自由度

となつて、段階推定の方がデータとの適合度がよいことがわかる。認知的不協和の存在はほぼ疑いのないことであるが、それを考慮したモデルの方がデータとの適合が悪いということは、そのモデルにおける取り込み方に問題があり、認知的不協和の影響が十分除去されていないと考えられる。しかし、選択モデルの推定結果を表2-3と表2-5で比較すると、適合度は、わずかにモデル1の方が優れているが、パラメーター各々についてみると、モデル2の方がその絶対値が大きくランダム効用の分散が小さくなっていると考えられ、モデル2の説明力はモデル1と同等程度にはあると考えられる。

また、推定法に根本的相違のある同時推定と、段階推定の比較を行なうために、モデル1とモデル2を比較する。そのためにモデル3の同時推定したパラメーターで得られた結果を段階的に検定し、その結果をモデル1の検定結果と比較した。

その結果は

LRS（model1）=       66.7       （39）  
LRS（model3）=       123.7      （39）（）内は自由度

となつて、LISRELモデル部分のデータとの適合度はモデル1の方がよい。これは、LISRELモデルだけで最尤推定を行なっているわけであるから当然の結果である。しかし、パラメーターを個別にみると同時推定の方が一般にt値が高いことから同時推定の推定値の方が信頼性が高いと考えられる。これは選択結果という情報を推定に同時に用いることにより、段階推定の改善の方向として述べた「認知的不協和」によるバイアスがある程度除去されたためと思われる。

また選択モデル部分の比較には、モデル3のLISRELモデル部分の推定値から、段階推定と同様に潜在変数の推計値（fitted value）を計算し、それと選択モデル部分の推定結果から、選択モデルの尤度を計算することにより、 $\overline{\rho^2}$ 値を直接比較した。その結果同時推定の $\overline{\rho^2}$ 値は0.173となり、段階推定の $\overline{\rho^2}$ 値0.352と比較するまでもなく適合の悪さを示した。しかし、同時推定の

表2-10 同時推定のLISRELモデルによる潜在変数を用いた選択モデルの推定結果

| 変数名                 | 潜在変数を用いたモデル    | 潜在変数を用いないモデル   |
|---------------------|----------------|----------------|
| 鉄道定数                | 0.121 (0.4)    | 0.538 (2.0)    |
| 費用                  | -0.0353 (-4.0) | -0.0268 (-4.2) |
| 幹線旅行時間              | 0.0461 (0.1)   | -0.405 (-1.6)  |
| 端末旅行時間              | -0.881 (-1.8)  | -1.57 (-4.2)   |
| 乗換回数                | -0.249 (-1.3)  | -0.195 (-1.3)  |
| ビジネスダミー             | 1.42 (3.6)     | 0.942 (3.4)    |
| 女性ダミー               | 0.662 (2.6)    | 0.466 (2.3)    |
| 快適性*                | 0.397 (1.9)    |                |
| 利便性*                | 1.05 (4.4)     |                |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.369          | 0.242          |

※LISRELモデルとあわせた全体の尤度比       （）内はt値

結果から求められた潜在変数の推計値の説明力を確認するために、その推計値を用いて選択モデル部だけ再推定した結果を表2-10に示す。その結果 $\overline{\rho^2}$ 値は0.369となり、モデル3から計算される潜在変数自体には大きな説明力があることが確認された。この二つの矛盾する結果の原因として、同時推定の選択確率の誘導における仮定とこの比較のための仮定が異なることと、選択モデル部分の定数項の推定値に問題があることが考えられる。表2-7と表2-10の推定結果の中で大きく異なるものは、定数項と幹線旅行時間、およびComfort\*の係数である。これらの中で特に問題となるのが定数項であつて、これは定数項と言うものがそもそも各個人の誤差項の平均値を表している。つまり残りの説明変数が全て等しいときどちらを取る確率が大きいかということであつて、各個人の好みが反映されているため、その値が個人によって正・負にばらつきがあり、それが統計的有意性を失わせる原因となり、モデル3の場合推定した結果たまたま大きな負の値をとり、これが適合度に影響を与えたと考えられる。

以上の結果から、同時推定と段階推定の優劣は一概に決めることはできないが、予測に用いる際の潜在変数の有意性が高い同時推定の方が、実用的であると考えられる。



また上記とは別に、潜在変数の推計値は、構造方程式、測定方程式のそれぞれの推定結果から、下の式によって別々に計算が可能である。

$$\text{構造方程式} \quad \mathbf{x}^* = \mathbf{B} \mathbf{s} \quad (2.63)$$

$$\text{測定方程式} \quad \mathbf{w}^* = \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} \mathbf{Y} \quad (2.64)$$

このなかで、特に構造方程式からのみ計算する意義は、このモデルで将来予測を行うとき、心理データの将来値を取ることは不可能なため、構造方程式のみ、つまり観測可能な変数だけを用いて定性的要因の将来値を計算できることである。

構造方程式と測定方程式から別々に定性的変数を計算したときの離散型選択モデルの推定結果はモデル1について表2-11、モデル2について表2-12に示してある。これを見ると、モデル1、モデル2ともに構造方程式から計算された潜在変数を用いた選択モデルの推定結果は、潜在変数の係数が、推定値、t値ともに小さくなり、潜在変数を入れないモデルと適合度はほとんど変わっていない。一方、測定方程式から計算された潜在変数を用いて推定さ

表2-11 段階推定の構造・測定方程式それぞれから推定された  
潜在変数を用いた選択モデルの推定結果

| 変数名                 | 構造方程式による潜在変数   | 測定方程式による潜在変数   |
|---------------------|----------------|----------------|
| 鉄道定数                | 0.533 (1.6)    | 0.415 (1.3)    |
| 費用                  | -0.0275 (-4.2) | -0.0348 (-4.1) |
| 幹線旅行時間              | -0.246 (-0.5)  | -0.0126 (-0.0) |
| 端末旅行時間              | -1.53 (-2.7)   | -1.38 (-3.0)   |
| 乗換回数                | -0.180 (-1.1)  | -0.332 (-1.8)  |
| ビジネスダミー             | 0.874 (2.8)    | 1.39 (3.7)     |
| 女性ダミー               | 0.472 (2.3)    | 0.662 (2.6)    |
| 快適性*                | 0.573 (0.5)    | 0.763 (2.2)    |
| 利便性*                | 0.0631 (0.1)   | 1.55 (4.1)     |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.230          | 0.356          |

( ) 内は t 値

表2-12 認知的不協和を考慮した段階推定の構造・測定方程式それぞれから  
推定された潜在変数を用いた選択モデルの推定結果

| 変数名                 | 潜在変数を用いたモデル    | 潜在変数を用いないモデル   |
|---------------------|----------------|----------------|
| 鉄道定数                | 0.544 (1.9)    | 0.418 (1.3)    |
| 費用                  | -0.0273 (-4.2) | -0.0339 (-4.1) |
| 幹線旅行時間              | -0.245 (0.3)   | -0.0199 (-0.1) |
| 端末旅行時間              | -1.56 (-3.2)   | -1.37 (-3.1)   |
| 乗換回数                | -0.187 (-1.2)  | -0.322 (-1.8)  |
| ビジネスダミー             | 0.904 (3.0)    | 1.34 (3.7)     |
| 女性ダミー               | 0.469 (2.3)    | 0.646 (2.6)    |
| 快適性*                | 0.497 (0.2)    | 0.914 (2.5)    |
| 利便性*                | 0.0230 (0.0)   | 1.54 (3.9)     |
| サンプル数               | 219            | 219            |
| $\overline{\rho^2}$ | 0.229          | 0.350          |

( ) 内は t 値

れた選択モデルの結果をみると、逆に構造方程式・測定方程式両方から計算された潜在変数を用いたモデルより適合度が高くなっている。

これら2つの結果から、潜在変数の構造全体を考慮して潜在変数の推計値を与える(2.25)式中で、構造方程式の影響を示す $\mathbf{B} \mathbf{s}$ による潜在変数の変動は小さいと考えられる。これは、LISRELモデルの構造方程式の説明変数の多くは、離散型選択モデルの効用関数の説明変数でもあるため、構造方程式が、離散型選択モデルの効用関数に新たに加える情報量は小さいためと考えられる。また、測定方程式からのみ計算される定性的変数に、大きな説明力があることは、やはり、どちらのモデルも認知的不協和の影響が大きいと考えられる。

## 2-6 本章のまとめ

本章では、潜在的要因を交通機関選択モデルに取り入れるための方法論を展開し、その有効性を確認するため、事例研究を通して実証的分析を行った。

本章で提案した方法は、図 2-1 に示されている消費者意思決定構造に基づいて、観測可能な変数から潜在的要因を推定する LISREL モデルと、潜在変数を含む離散型選択モデルについて段階推定と同時推定の 2 種類の方法を提示し、それぞれ改良の方向性を探るモデルを提案した。この中で、アンケート調査から得られた知覚値や態度といった潜在変数の数理心理学的指標は、LISREL モデルの、観測可能な変数として用いられている。このことは、2-1 で述べた過去の研究例のように、単に数理心理学的指標を、交通機関選択モデルの効用関数の説明変数に加えるものと異なり、心理指標と潜在変数の関係を測定方程式で、潜在変数と定量化可能な変数の関係を構造方程式で、それぞれ表しているため、効用関数の説明変数として、構造方程式だけから計算される潜在変数を用いることが可能なことが特長である。つまり、本モデルを将来予測へ適用するときに、心理指標の将来値は不要であるため実用性が高いと思われる。

2-5 で示した実証的研究の結果を見ると、段階的に LISREL モデルより計算された 2 つの潜在変数は、選択モデルにおいて 2 つとも大きな説明力を有しており、潜在変数を含まないモデルと比較すると、モデルの適合度も大幅に良くなっていることがわかる。ただし、測定方程式から計算された潜在変数を使用したモデルと、構造方程式から計算された潜在変数を使用したモデルを比較すると、測定方程式から計算された潜在変数に大きな説明力があるということがわかった。本モデルを将来予測に適用する際、構造方程式から計算された潜在変数を用いるので、潜在変数の将来値の有意性に問題がある。この原因として、本論文で行われた事例研究では、既存のデータを用いたため、データ項目の制約によるミスマスペシフィケーションが大きな原因のひとつであると考えられる。例えば、意思決定者の心理的要因を決定する重要な指標であると思われる個人の収入、社会的地位などがデータに無く、それらが説明変数として用いられなかったことがモデルの適合を悪くしていたと考えられる。また、全システムの同時推定は、その統計的性質や、有意性に優れているが、それを将来予測に用いる際に重要な、推定結果中の線形構造方程式モデル部分から計算される推計値を用いて、選択モデルを推定した場合に、選択モデルの適合度が段階推定よりも低いという問題があった。これは、同時推定モデルの選択モデル定数項の推定値が特異な値になったためと考えられるがその原因として、選択したモードの相違による主観的価値基準の違いは無視できない（河上、広島（1985））が、本研究ではそのことを考慮していないためにそのことが顕著に現われる定数項の推定値が特異な値になったと考えられる。また、知覚値指標の離散性を考慮したモデルは、他のモデル

との比較のため同じ特定化を行なったが、モデルの定式化をデータ本来の性質に近づけたモデルであるので、別の特定化を行なうことにより、より適合度が高く説明力のあるモデルになることが可能であると考えられる。

以上本研究の結論として、提案した各モデルは潜在変数を用いないモデルより適合度や説明力の点において優れた結果を示しているが、それぞれのモデルの優劣については、一長一短があり一概に言うことはできなかった。

本研究の今後の課題としては、推定の手法に関して比較検討するための理論的基礎がしっかりとできていなかったため、比較があいまいに終わってしまった。また知覚値指標の離散性を考慮したモデルは、その特定化の変更により解釈の容易な説明力の高いモデルになることが期待されるが、それを十分行なっていない。これらのことをふまえて異なるデータによるさらなる実証的研究が、この手法の有効性を確認するために必要であろう。

## 第2章 参考文献

青山吉隆, 芝原靖典, 岩瀬広, 村上幸二郎 (1988) 交通機関分担モデルにおける影響要因選定の分類と構造に関する研究, 土木計画学研究・論文集, No.6, pp.193-200.

秋山学 (1995) 消費者行動, 蓮花一己・西川正之編, 現代都市の行動学 15 章, 福村出版, pp.151-160.

太田勝敏, 原田昇 (1982) 非集計行動モデルの研究の現状と課題, 土木計画学研究・講演集, No.4, pp.375-384.

河上省吾, 広島康裕 (1985) 利用者の主観的評価を考慮した非集計手段選択モデル, 土木学会論文集, No.353/IV-2, pp.83-92.

鈴木聡, 原田昇, 太田勝敏 (1986) 意識データを用いた非集計モデルの改良に関する分析, 土木計画学研究・論文集, No.4, pp.229-236.

林知己夫, 駒沢勉 (1982) 数量化理論とデータ処理, 朝倉書店.

原田昇, 兵藤哲朗 (1991) 交通行動分析のフロンティアー混乱と仕切り, 土木計画学研究・講演集, NO.14(2), pp.9-16.

松田紀之 (1988) 質的情報の多変量解析, 朝倉書店.

森川高行 (1990), ステイティッド・ブリファレンス・データの交通需要予測モデルへの適用に関する整理と展望, 土木学会論文集, No.413/IV-12, pp.9-18.

森地茂 (1984) 非集計行動モデルによる予測, 土木計画学講習会テキスト 15: 非集計行動モデルの理論と実際, 土木学会, pp.121-147.

屋井鉄雄, 森川高行 (1991) 交通需要モデル研究のダイナミズムー10年の軌跡, 土木計画学研究・講演集, NO.14(2), pp.1-8.

柳井晴夫, 繁樺算夫, 前川眞一, 市川雅教 (1990), 統計ライブラリー因子分析ーその理論と方法ー, 朝倉書店.

山田菊子 (1991), 系列相関を持つ RP データと SP データを同時に用いる交通行動モデルの推定法, 京都大学大学院工学研究科応用システム科学専攻修士論文.

渡辺隆, 岩崎征人, 杉本巧 (1987) 新交通システム導入計画のための需要予測モデルの作成, 土木計画学研究・講演集, No.10, pp.85-92.

Amemiya, T. (1985) Advanced Econometrics, Harvard University Press.

Assael, H. (1987) Consumer Behavior and Marketing Action, 3rd ed., Kent.

Balace, C.R.M. (1988) Improvement of Data Collection and Modeling Techniques of Disaggregate Logit Models, Ph.D. Dissertation, Department of Social Engineering, University of Tsukuba.

Ben-Akiva, M. and Lerman, S. R. (1985) Discrete Choice Analysis, Theory and Application to Travel Demand, MIT Press.

Ben-Akiva, M. and Morikawa, T. (1990) Comparing Ridership Attraction of Rail and Bus by Estimating Mode choice Model, Prepared for Presentation at the 69 th. Annual Meeting of T.R.B., Washington D.C..

Dobson, R. and Tischer, M. L. (1977) Comparative Analysis of Determinants of Modal Choice by Central

Business District Workers, Transportation Research Record, No.649, pp.7-14.

Golob, T.F., Horowitz, A. D. and Wach, M. (1979) Attitude Behavior Relationships in Travel Demand Modeling, Behavioral Travel Modeling, Henscher and Stopher (eds.), London, Croom Helm, pp.739-757.

Joreskog, K. and Sorbom, D. (1984) LISREL VI Analysis of Linear Structural Relations by Maximum Likelihood, Instrumental Variables, and Least Square methods, User's Guide, Department of Statistics, University of Uppsala, Sweden.

Koppelman, F. S. and Pas, E.I. (1980) Travel-Choice Behavior: Models of Perceptions, Feelings, Preference, and Choice, Transportation Research Record No.765, pp.26-33.

Koppelman, F. S. and Patrica, K. L. (1981) Attitudinal Analysis of Work/School Travel, transportation Science, Vol.15, No.3, pp.233-254.

Levin, I. P. (1979) The Development of Attitudinal Modeling Approaches in Transport Research, Behavioral Travel Modeling, Henscher and Stopher (eds.), London, Croom Helm, pp.758-781.

Louviere, J.J. (1979) Attitudes, Attitudinal Measurement and the Relationship between Attitude and Behavior, Behavioral Travel Modeling, Henscher and Stopher (eds.), London, Croom Helm, pp.782-794.

Lovell, C.H. (1975) Modeling the Modal Choice Decision Process, Transportation, 4, pp.253-265.

McFadden, D. (1986) The Choice Theory Approach to Market Research, Marketing Science, Vol.5, No.4, Fall, pp.275-297.

Morikawa, T. (1989) Incorporating Stated Preference Data in Travel Demand Analysis, Ph. D. Dissertation, Department of Civil Engineering, MIT.

Morikawa, T., Ben-Akiva, M. and McFadden, D. (1990) Incorporating Psychometric Data in Econometric Travel Demand Models, prepared for Banff Invitational Symposium on Consumer Decision Making and Choice Behavior.

Recker, W. W. and Golob, T.F. (1976) Attitudinal Modal Choice Model, Transportation Research, Vol.20-A, pp.293-310.

### 第3章 個人の異質性を考慮した交通行動分析

#### 3-1 はじめに

本章では、個人の異質性を考慮した非集計分析の方法論を展開し、S P、R Pおよび態度データを用いて意思決定者の嗜好の違いを考慮した交通手段選択モデルの提案を目的とする。

非集計モデルを用いて個人の嗜好の違いを考慮する手法は一般に次の3種類に分類される(片平(1987-1))。1つは個人属性別やトリップ目的別などの外的属性を用いる方法である。その中には、あらかじめ基準を定めて、母集団をいくつかのセグメントに分けるアプリアリセグメントを行う方法と、効用関数の説明変数に外的属性を導入することで、嗜好のばらつきを表現する方法の2つがある。2つめは内的属性、つまり効用や態度などの多次元的に定められた変数を基準として、それらが同質なグループにサンプルを分類するセグメンテーションである。具体的には、個人ごとにロジットモデルを推定して得られた個人ごとの効用パラメータと、いった、意思決定者の嗜好を表す変数による母集団の分割を指す。3つめはランダム係数モデルなど非集計モデルのそのものの構造を異質性を取り込んだかたちで定式化し直すモデルの開発である。これらの中で第1の手法である外的属性を用いてセグメンテーションを行う方法は容易でかつ解釈がしやすいなどの利点を持ち、セグメンテーションの意識をせずに使われることが多い。しかし、外的属性によって分けられたサブグループが同質であるという保証を得ることは難しい。2番目に示された、嗜好が同質な集団に分割することを目的とするベネフィットセグメンテーション手法は、これまでマーケティングの分野では1960年代から研究が進んできているが、交通行動の分野での適用例は少なかった。これは主に、一般的な消費財の購買行動を対象とするマーケティングリサーチでは、主婦層を対象にした洗剤のブランド選択など、対象とする個人属性や購入目的などが同質の個人をターゲットとして行われる場合が多く、母集団に属する人の属性や選択肢の属性が似ている場合が多い。このため、分析対象となる選択行動の嗜好のばらつきは、個人の嗜好の違いが主要な要因となり、個人の異質性を考慮した選択行動分析が盛んに行われてきたのである。交通行動分析が対象とする交通行動の分野、特に非集計モデルが分析の道具としてよく用いられている交通機関選択分析では、一般に対象となる選択肢の属性や交通目的などが個人間で異なっている場合が多く、個人の異質性の影響に比べて選択肢間の属性の値の差異や交通目的などの外生的要因が選択に大きな影響を与える場



合が多い。例えば、目的地が与件である場合には選択可能な交通機関が限定されることや、自由目的トリップと業務トリップでは時間価値やトリップチェインパターンが異なることなどが挙げられ、一般によく行われる交通目的別のアプリオリセグメンテーションは、個人間の差異ではなく、選択の外生的要因によってパラメータが変動する影響を考慮したものである。このようなことより、交通機関選択は選択に大きな影響を与える外的属性によるセグメンテーションが、有効であると考えられているが、近年観光交通や休日交通などの非日常交通が交通計画上の問題として取り上げられるようになり、その分析が盛んになってきた。一般に非日常交通は通勤、通学、業務交通と比較して交通制約が緩く、その行動には嗜好の違いが顕著に現れてくると考えらる。そのため、嗜好の違いを考慮したベネフィットセグメンテーションの有効性が高い。また、日常交通においても、交通サービスや交通政策は鉄道会社間の時間短縮競争、快適性車両の導入、相互乗り入れ、あるいは渋滞（所要時間）・駐車場に関する情報提供等に代表されるように多種多様化しており、そのような政策に対してはより詳細な需要予測が必要になってきている。そのような多様な交通サービスの中から交通選択を行う場合、所要時間を重視する人、快適性を重視する人などの異なる嗜好を持った人たちはそれぞれ同じ代替案属性に対して異なる反応を示し、交通サービスに対するニーズや評価は明らかに母集団内で均質ではない。したがって、ある交通サービスを導入した場合、そのサービスの変化により得られる効果を測定するためには、その反応の度合いの違いを考慮した分析が必要となる。このように交通サービスの多様化に伴って、個人の嗜好の違いは交通行動分析において無視できないものとなってきており、個人の異質性を考慮した分析の必要性が高まってきていると言える。そこで異質性を分析するにあたっては、これまで数多くの研究例があるマーケティングリサーチの分野で行われてきた成果を参考に、交通行動分析、特に交通機関選択に焦点を絞り分析手法をいくつか提案し、その有効性を検討するものである。

また、このような異質性の分析とは別に、仮想の状況に対する選好意思表示であるS Pデータの研究も盛んに行われている。S Pデータは実際の行動結果であるR Pデータと異なり、条件の設定方法によって制約条件の影響を小さくすることができ、属性間のトレードオフの影響が相対的に高まることや、同一個人から多数の情報を得ることが容易であり、嗜好の違い、つまり効用パラメータの個人による違いを分析することに適していると考えられる。このような特性を持つS Pデータを個人の異質性を考慮した分析に適用することは非常に効果的であると

表3-1 R PデータとS Pデータの比較（斜体部原本より変更）

|        | R P データ         | S P データ            |
|--------|-----------------|--------------------|
| 選好情報   | 実際の行動結果に基づく     | 仮想の状況における意思表示      |
|        | 市場における行動と一致     | 市場での行動と不一致の可能性     |
|        | 得られる情報は「選択結果」   | 「順位付け」「評点づけ」「選択」など |
| 代替案    | 現存しない代替案は取り扱えない | 現存しない代替案も取り扱える     |
| 選択肢の属性 | 定量的属性のみ         | 定量的および定性的属性        |
|        | 測定誤差があることが多い    | 測定誤差はないが知覚誤差の可能性   |
|        | 属性値の範囲が限られている   | 属性値の範囲を拡張できる       |
|        | 属性間の重共線性が大きい    | 属性間の相関を制御できる       |
| 選択肢集合  | 不明瞭             | 明瞭                 |

考えられ、その有効な分析手法の確立はS P分析が効果的である局面の存在を確認する意味でも重要である。S Pデータの詳しい性質については森川（1991）にまとめられている。それによると、S Pデータの最大の問題点はその信頼性であり、信頼性はS Pデータの選好と市場での選択の一致性を指す信憑性と、S P実験の条件による回答の変動を指す安定性の2面を持つとしている。森川（1991）はR PデータとS Pデータの特質を表3-1のようにまとめている。

このようにS PデータとR Pデータは相互補完的關係にあり、これらを同時に用いたモデルが、Ben-Akiva and Morikawa（1990）によって提案されている。また、その発展形として誤差項の系列相関を考慮したモデルをMorikawa, Ben-Akiva and Yamada（1992）が提案し、系列相関によって生じるパラメータのバイアスを修正している。これらのモデルについては後で詳しく説明する。

3-2 個人の異質性を考慮した既存研究のレビュー

これまで行われてきた非集計モデルを用いた交通需要予測分析では、対象母集団を均質とみなして、得られたサンプルが同一の効用パラメータを持つと仮定してモデル構築を行うか、異

表3-2 アプリオリセグメントに用いる主要な基準（田中（1992））

| 基 準           |                        |                               |
|---------------|------------------------|-------------------------------|
| 意思決定者の特性による基準 | 地理的特性（Geographic）      | 市町村<br>郊外／都心<br>気候            |
|               | 人口統計的特性（Demographic）   | 性別<br>年齢<br>家族人数<br>ライフステージ   |
|               | 社会経済的特性（Socioeconomic） | 年収<br>職業<br>宗教<br>教育程度        |
|               | 心理的特性（Psychographic）   | ライフスタイル<br>パーソナリティ            |
| 意思決定者の反応による基準 | 行動科学的（Behavioral）      | 購入量<br>ロイヤルティ<br>態度<br>ベネフィット |

質性を考慮する場合は個人の性別や職業、収入といった社会経済属性やトリップ目的などのトリップ属性によってアプリオリにセグメントを規定し、いくつかの同一属性のセグメントに母集団を分割して、それぞれのセグメントごとに異なるモデル構築を行うことが多い。表3-2に主なアプリオリセグメントの基準を示す。

これまで交通行動の分野では、交通目的以外にも、表3-2中に示される基準の中で、主に地理的特性や、人口統計的特性、社会経済的特性によってセグメンテーションされることが多く、心理的特性や行動科学的基準に基づいてアプリオリセグメントが行われた例は少ない。ここで、個人属性などの意思決定者の特性によるセグメンテーションは、セグメント内の各個人が同質であるということに対する根拠に乏しく説得力に欠ける場合もあり、非集計行動モデルの基礎が確立された直後から、外的属性によるセグメンテーションに代わるアプローチから個人の異質性を考慮しようと試みている研究がなされてきた。そして、それらは表3-2でいう

求める便益（ベネフィット）もしくは2章でも述べた態度を基準とした、つまり効用パラメータの同質性を基準としたセグメンテーションによるアプローチと、ランダム係数モデルのような、ロジットモデルに代表される離散型選択モデルの再定式化型のアプローチに大別できることを3-1で概説した。ベネフィットセグメンテーションによるアプローチは先の外的属性によるアプローチと同じく、同一セグメントでは同一の効用パラメータを持つことを仮定しているが、より効用関数の同質が保証されうるセグメンテーション指標の探求を目指すものである。一方、ランダム係数型アプローチでは母集団内で異なる効用パラメータの分布を、ある特定の確率分布を仮定したり、ノンパラメトリックな分布を仮定することによって異質な母集団全体に同一のモデルを適用することを目指すものである。以下ではそれぞれのアプローチに関する既存の研究を整理する。

片平（1987-2）によると、マーケットセグメンテーションを行う際に注意すべきこととして、次の3点を上げている。

- 1）形成されるセグメントは基準に関して十分に同質的でなければならない。
- 2）セグメントはそれを形成するのに十分な大きさがなければならない。
- 3）各セグメントは他と区別して接近可能でなければならない。

効用パラメータの同質性に基礎をおくベネフィットセグメンテーションで特に重要となるのが、1）に挙げられたいかにして何らかの意味で同質な集団に母集団を分割するかということである。つまり個人の異質性を考慮したモデルの構築には同質の効用パラメータを持つ個人どうしでセグメントが形成されることが望ましいため、個人ごとにパラメータ値が求められる場合にはそれを基準にしてセグメンテーションを行うことが最も望ましいと考えられる。その考えに基づいて、個人パラメータを測定し、セグメンテーションを行った研究としては Currim（1981）、Ogawa（1987）が挙げられる。個人モデルに関する研究としてはマーケティングの分野ではコンジョイント測定法と呼ばれる手法が主流になっている。コンジョイント分析とは数理心理学におけるコンジョイント測定法の考え方をマーケティングに応用しようという試みを総称したものである（Louviere（1994））。ここで、コンジョイント測定法とは大澤ら（1980）や小川（1981）によると「多次元的な要因の組み合わせに対する何らかの順序関係が与えられ

たときに、そこから個々の要因の効果を測定する個別尺度、及び与えられた順序関係を再現する総ての要因の同時結合尺度をある定められた結合ルールのもとで同時に見いだすこと」である。具体的には、コンジョイント分析法として代表的なアルゴリズムとして MANOVA (Kruskal (1965)) や LINMAP (Srinivasan and Shocker (1973)) があげられるが、これらはいずれも確定論的考え方に基づくものであり、確率論的考え方に基づくものとして RANKLOGIT (小川 (1981) など) があげられる。交通計画の分野では高田・湯沢 (1988) が、コンジョイント分析にロジットモデルを導入した、コンジョイントロジットモデルと呼ばれる確率論的手法を用いて交通機関選択を分析している。また、伊藤・湯沢 (1990) は、そのコンジョイントロジットモデルを企業立地選択に適用して、個人(企業)別のパラメータをクラスター分析や割付法を用いてセグメンテーションを行い、異質性を考慮した分析を行っている。また森川・白水 (1991) は一対比較質問の5段階評定のデータを用い重回帰分析によって個人パラメータの近似値を求め、その値に基づいてクラスター分析や視覚的分割法によってセグメンテーションを行い、それらのセグメンテーションの有効性を実証している。このようなアプローチには、個人パラメータ推定のための同一個人からの観測数の確保の問題や、個人パラメータの信頼性の問題、クラスタリング手法等の問題点が存在するが、Ogawa (1987) は効用パラメータの推定とグループ分けを同時に行う手法を提案し、それらの問題点を解決しようと試みている。片平 (1987) は Wishart (1969) の多峰型のヒストグラムをいくつかの単峰型に分割するモード分析手法を用いて、これらの問題を回避することのできるモデルを提案している。

このような手法とは別に、非集計モデルの基本的定式化の修正を行うものとして、ランダム係数型アプローチがある。ロジットモデルをはじめとして通常、交通需要予測に用いる非集計モデルは、対象母集団の効用パラメータは固定されており固定係数モデルとも呼ばれるが、ランダム係数モデルは効用パラメータが個人間で何らかの分布に従う確率変数であるとするものである。このようなランダム係数モデルはロジットモデルに限らず、回帰モデルなどでも研究が行われている (Nicholas (1993))。このモデルは個人間でパラメータは同一と仮定するよりも理論的には優れており、ロジットモデル開発の初期から提案されている。Daganzo (1979) はプロビットモデルにおいて、パラメータに正規分布を仮定することによって個人の異質性を考慮することができるということを論じている。このモデルは、簡単な再定式化を行うことによって一般化プロビットモデルの定式化に帰着する。Fisher and Nagin (1981) は仮想の駐車場の2

項選択を例に一般化プロビットモデルの有効性に関する実証的分析を行っている。しかし、この再定式化された一般化プロビットモデルは3項以上の多項選択においてはその推定が非常に困難となり、その推定法に関する最近の研究としては McFadden (1989)、Bolduc and Ben-Akiva (1990)、Lam and Mahmassani (1990) 等が挙げられる。その他にもマーケティングの分野ではガンマ分布を仮定するモデル (例えば Gupta (1991) など) が用いられることもある。このように効用パラメータ総てにパラメトリックな分布を仮定するモデルではなく、一部の変数をパラメトリックに分布させる **Mixing Distribution Model** やノンパラメトリックに分布させる **Mass Point Model** などが西井ら (1995) によって目的地選択モデルに適用されている。Mixing Distribution Model をパネルデータに用いて個人の異質性を検討した研究に Hensher and Wrigley (1986) があり、Mass Point Model を用いてパネルデータの消耗バイアスに存在する個人の異質性を検討したものに藤原ら (1996) がある。これらのパネルデータを用いた異質性の研究については4章で解説する。Mass Point Model の一般型である潜在クラス分析の考え方を応用した潜在セグメントモデルが、近年マーケティングリサーチの分野で適用例が多く見られる (例えば Kamakura et. al (1994))。これは各個人を観測不能な真のセグメントに分類するのは不可能であると考え、各個人は各セグメントに確率的に所属するものとして、各個人のセグメントへの所属確率とセグメントごとのパラメータ値を同時に推定するものである。潜在セグメントを考慮したモデルは他に Kamakura and Russel (1989)、片平 (1991)、Swait (1993)、Kamakura et. al. (1995)、Kamakura and Wedel (1995)、Gopinath and Ben-Akiva (1995) らがある。この手法については後に詳しく解説する。また、離散型選択モデル以外には、Bhat (1996) が帰宅トリップでの買物行動を生存時間モデルを用いて分析し、その中で個人の観測不能な異質性の問題を取り扱っている。

### 3-3 個人に繰り返し尋ねた S P データを用いた個人ベースの行動モデル

#### 3-3-1 個人から繰り返し採取した S P データの性質について

S P データは分析者が自由に条件などを設定できるため、個人から容易に複数回の回答を得ることができる。このことは、同一個人からより多くの情報を引き出すことができるため、個人の行動特性の把握が容易になり、異質性を考慮したモデル構築に用いるために有効であると



考えられる。しかし、複数S Pデータは4章で詳しく述べるパネルデータと同様に、3-1で述べたS P固有の問題だけでなく他にもいくつかのバイアスを持っている可能性が指摘できる。ここでもう一度森川（1990）に基づいてS Pデータ固有の問題である信頼性について詳しく述べる。S Pデータの信頼性とは大きく分けて2つに分類される。一つはS Pデータの信憑性であり、もう一つはS Pデータの安定性である。S Pデータの信憑性とは、S Pデータは選択結果が被験者の効用レベルに影響を与えないため、実際の行動の場合とは異なった意思決定のプロセスによって答えてしまう問題であり、S Pデータの安定性とは、S Pデータの属性や選択肢の設定の仕方によって回答が変動する問題である。このような特性を持つS Pデータを同一個人から複数採取することによって、以下のような影響が加わると考えられる。

- 1) 被験者の疲労によるデータ精度の低下 (Fatigue effect)
- 2) 複数回の回答が意思決定に影響を与えそれぞれの選択が独立性を失う (Inertia effect)
- 3) 最初の方の質問が練習となって信頼性が落ちる (Warming up effect)

具体的には、1) は複数の質問を行うことにより、被験者は疲労して回答の回数が増えるに従いその回答にいい加減さ（ノイズ）が増してゆくことで、2) は惰性で同じ選択肢を選び続けたり、2章で述べた認知的不協和解消行動の問題と同様の問題が生じたりする。ただし、この惰性によって同じ選択肢を選択し続けるのかということは、設定条件以外の要因で選択が行われている場合に観測上同一となるため、分析者の側から判断することは困難である。3) は1) と逆で、複数回答が可能なため、質問の設定を理解する前に回答を行ってしまい、最初の回答の信頼性が低い問題である。これらの問題について、推定したモデルの推定値および標準誤差を評価指標として分析した事例がいくつか報告されている。例えばBen-Akiva, Morikawa and Shiroishi (1991) がS Pの順位付けデータを複数回答とみなしてパラメータの安定性を分析し、異なる順位ではパラメータが有意に異なると報告している。また藤原（1993）は3回繰り返したS P質問を分析し、3回までのS P質問では安定性や信頼性に問題は生じないと言う結論を出している。複数S Pデータの信頼性については、条件設定などにも依存するため一概に何回までならば信頼できるということは言えないが、質問回数や与える条件の複雑さなどの調査の設計を、個人から得ることのできる情報量とノイズ量の間関係を見極めた上で行わなければならない。

### 3-3-2 複数S Pデータによる個人パラメータの推定

個人ごとのロジットモデルによって推定された個人パラメータを指標としてセグメンテーションを行う場合には、パラメータ値の扱い方に関して次の3点について留意しなければならない。

- ・ロジットモデルによって推定されたパラメータは、誤差項のスケールパラメータとの積の値で得られるということである。同一のパラメータを持つ個人でも誤差項の分散が小さい場合は、推定されるパラメータの値は相対的に大きくなり、分散が大きい個人は逆に相対的に小さくなる。ベネフィットセグメンテーションを行う場合、個人の異質性は各要因間のトレードオフの大小関係によって捉えるのが本来的であるので、誤差項の影響を除去する必要がある。
- ・各説明変数の測定単位にパラメータのスケールが依存しているという問題である。つまり個人ごとのパラメータ値の分布が、パラメータにかかる変数が時間なのかあるいは費用なのかといった尺度の違いや、それぞれの説明変数の単位の取り方に依存しているため、恣意的にセグメントを操作できる可能性がある。
- ・ロジットモデル等の離散型モデルにおける小サンプルの問題（例えば伊藤、湯沢（1990）、森川、白水（1991））である。個人ごとのデータ数が限られているため、例えば最大尤度が一意に定まらなかったり、推定可能であったとしてもパラメータ推定値そのものの自体には厳密な意味で統計的推論が行えないということである。

以上の3点を考慮してパラメータ値をそのままセグメンテーションの指標とせず、一つのパラメータ値を基準としたパラメータ値の比を指標として採用するのが妥当であると考えられる。この考え方に基づくと、個人パラメータを用いたセグメンテーションによるアプローチは、図3-1に示すフローにしたがって行われる。



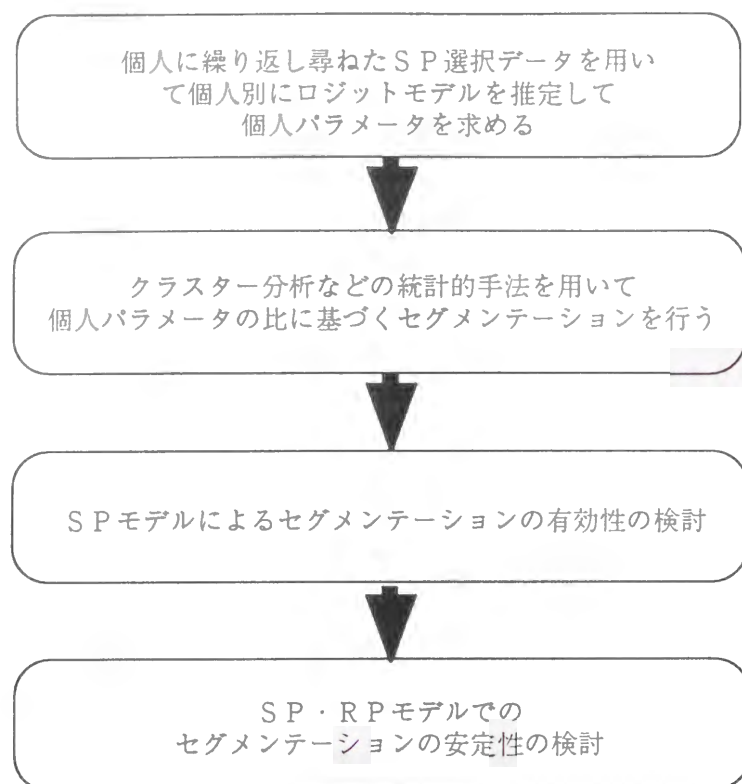


図3-1 個人パラメータアプローチによるセグメンテーションのフローチャート

### 3-3-3 個人パラメータに基づいたセグメンテーション

本研究で用いるデータは、首都圏の湾岸海浜観光地域のある半島を訪れていた人を対象に、無作為にアンケート調査票を配布して郵送回収されたものである。アンケートの内容としては、主に次の5種類に分類することができる。

- ・個人の社会経済属性に関するデータ
- ・性別、年齢、職業、自家用車の有無
- ・今回の旅行の手段、経路、旅行内容に関するデータ（RPデータ）
- ・目的、旅行人数、旅行ルート、旅行費用、交通機関選択理由（態度指標）等
- ・新規航路の選択に関するデータ（SPデータ）

その湾岸海浜観光地域と対岸を結ぶ仮想の新規高速船（フェリー）が就航したという仮定のもと、新規高速船の所要時間は一定とし、各被験者に6種類の料金と運行間隔の組み合わせを示し、新規高速船を利用するか否かの2項選択データを各被験者より6組得ている。なお、実際の利用交通手段が車である人と車以外（本研究においては鉄道のみ）である人では、それぞ

れ新規高速船へのアクセス・イグレスは現状の利用交通手段を変更しないという条件をつけ、異なる料金設定でSP質問を行った。

全サンプル852人のうち、今回の旅行と同条件（旅行人数、旅行目的が同じ）の場合に実際に新規高速船の利用が可能であると考えられ、上記の各データについて有効な回答が得られたサンプルは211人であった。ここで利用不可能な人とは、新規高速船の運賃・所要時間・運行頻度などの条件に関わらず利用しないと回答した人、当該地域在住の人、観光バス利用者で交通機関選択が不可能な人である。このうち、全ての人がSP質問6個に回答したわけではないのでRPデータ211、SPデータ819のデータが得られている。また、個人パラメータ推定に必要なデータが得られたサンプルは84人であった。

SP質問は新規高速船の運賃と運行頻度を政策変数として変動させているので、個人モデルに用いることのできる説明変数は以下の3つである。

- 1) 新規高速船定数項
- 2) 総費用（単位：1000円）
- 3) 新規高速船の運航頻度（単位：運航本数／時）

つまり、個人モデルは3係数を持つ効用関数を仮定し、6つの選択結果の観測値から最尤推定法を用いて推定を行っている。このような自由度の低いモデルから得られるパラメータは、統計的に有意であることが非常に難しいが、6つの選択結果から合成した個人指標のベクトルであると考え、マーケットセグメンテーションの指標として用いることは可能であると考えられる。

これらの説明変数を用いて、個人ごとに新規高速船を利用するか否かの2項ロジットモデルを推定した。「新規高速船を利用しない」と回答した場合の選択は実際に利用した交通手段・経路となる。なお84サンプル中、1人について個人パラメータの推定値が収束しなかった。これは3-3-1で述べた少数サンプルによる解の不定性の問題が生じていると考えられる。ここで、先に述べたように個人ごとに推定されたパラメータ値をそのままセグメンテーションの基準とするのではなく、あるパラメータを基準としたパラメータ比をセグメントの基準として用いる。本研究においては総費用の個人別パラメータがほぼ全サンプルを通じてマイナスの符号となったため、パラメータ比の値の符号が解釈しやすいことから、定数項および運航頻度のパ

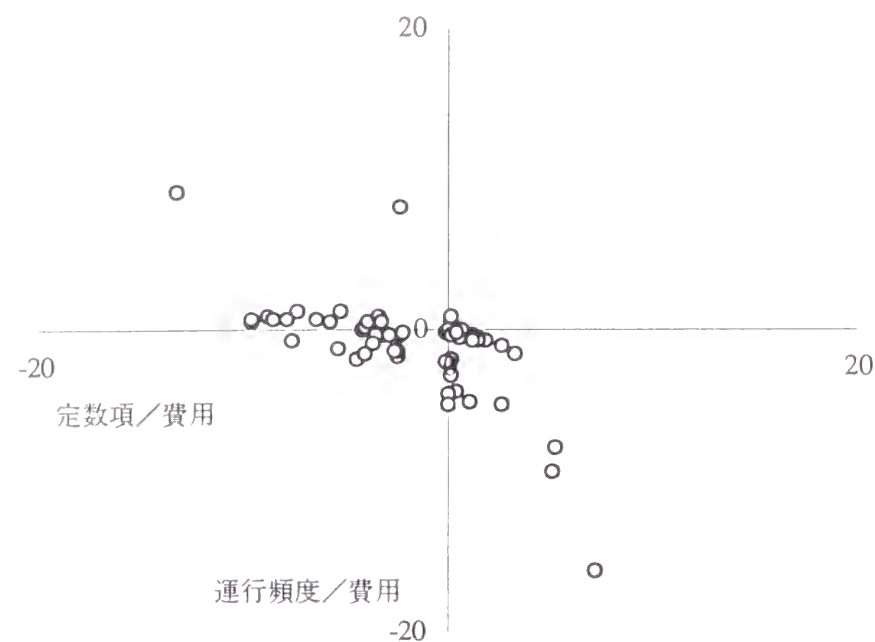


図 3-2 個人パラメータ比の散布図

ラメータ値を総費用のパラメータ値で除した値をセグメンテーションの基準とすることとした。このパラメータ比の値の散布図が図 3-2 である。この 2 次元のベクトルをもとにクラスター分析を行い、3-2-1 で述べたクラスタリングの基準に従い 2 つのセグメントを抽出した。各セグメントごとの散布状況を示したのが図 3-3 である。今回のクラスター分析を行うときの類似度の測度として平方ユークリッド距離を用いており、視覚的には座標平面上の各点間の距離の近いもの同士がセグメントを形成することになる。なお、個人パラメータ推定者 83 サンプル中の 1 人については、パラメータ値の比の値がかなり大きな値となったために図 3-2 および図 3-3 においては含まれていない。

図 3-3 から分かるように、ほぼ定数項/費用軸の正負の値によってセグメンテーションされる結果となっている。これは、大多数の人が運行頻度/費用軸近傍において散布しているためであり、総費用に対するパラメータ値がほぼ負の値を取っていたことから判断すると、セグメント 1 は定数項が負であるグループ、セグメント 2 は定数項が正であるグループに分割されたと見える。

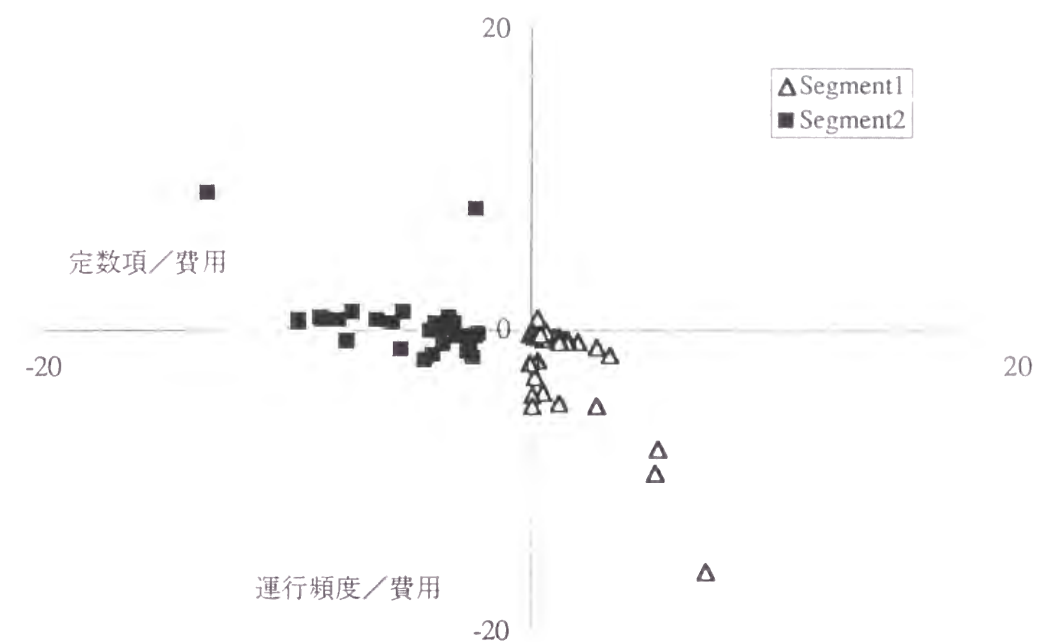


図 3-3 個人パラメータのクラスター分析結果

このようにして得られた各セグメントごとにデータをプールして個人モデルと同様な 2 項ロジットモデルによって、各セグメントごとの未知パラメータを推定した結果を表 3-3 に示す。なお、各セグメントともにデータをプールすることによって総所要時間に対するパラメータ値も推定可能となるので、個人パラメータ推定に用いた 3 つの説明変数に総所要時間（単位：10 分）を加えて 4 つの説明変数によってモデルの推定を行った。推定されたパラメータ値はどれも妥当な符号を持っており、個人パラメータの定数項の符号でセグメンテーションされたとおり、各セグメントは、ともに定数項にはっきりその違いが現れている。また、尤度比検定によって両セグメントのパラメータベクトルの等価性は有意水準 5% で棄却され、 $\bar{p}^2$  値もセグメンテーションを行わないモデルと比較して大きく向上している。ただし、時間に対するパラメータの t 値は両セグメントともに、特にセグメント 2 において有意な値とはならなかった。この理由としては、もともと総所要時間をあまり考慮していないということも考えられるが、SP 質問の設定において新規高速船の所要時間を一定であると仮定しているために、回答者の注意が変動させた運賃や運航頻度に向けられてしまったという可能性も高いと考えられる。費用に対するパラメータの絶対値はセグメント 2 の方がセグメント 1 よりも 2 倍大きく、逆に運航頻

表 3－3 個人パラメータを基準とするアプリオリセグメント別の推定結果

|                | 全データ             | セグメント 1       | セグメント 2                   |
|----------------|------------------|---------------|---------------------------|
| フェリー定数         | -1.69 (-6.8)     | -3.55 (-6.6)  | 1.28 (2.0)                |
| 総費用            | -0.312 (-7.4)    | -0.340 (-3.9) | -0.704 (-5.4)             |
| 運行頻度           | 1.91 (8.3)       | 4.02 (6.1)    | 1.06 (3.9)                |
| 総所要時間          | -0.0649 (-2.1)   | -0.103 (-1.7) | -0.0293 (-0.8)            |
| 観測数            | 492              | 240           | 252                       |
| $\rho^2$       | 0.270            | 0.381         | 0.318                     |
| $\bar{\rho}^2$ | 0.255            | 0.325         |                           |
| $\chi^2$ 値     | $\chi^2 = 42.69$ |               | $\chi_{0.05}^2(4) = 9.49$ |

( ) 内は t 値

度に対するパラメータの絶対値はセグメント 1の方がセグメント 2よりも 4 倍大きくなっている。これを単純に比較すれば、セグメント 1は S P 質問の運航頻度の変動により敏感に反応し、セグメント 2は運賃の変動により敏感に反応するセグメントであるといえる。セグメント 1に関しては、前述した総所要時間に対するパラメータの t 値は有意水準 10%では有意で、パラメータ値の絶対値もかなり大きい値となっており、運航頻度にもより敏感であることを考えあわせると時間に関する利便性をより重視するセグメントであるといえる。この結果から、新規高速船を利用するか否かという観点に立つと、セグメント 1は定数項が負であるので新規高速船の利用意向そのものは低いが、時間に関する利便性を増せばより効率的に高い利用率が見込めるセグメントであり、一方セグメント 2は、定数項が正であるので利用意向そのものも高く、安い運賃の設定によってよりいっそう高い利用率が見込めるセグメントであると推察できる。

ここで Ben-Akiva and Morikawa (1990) で提案されている S P データと R P データを同時に用いたロジットモデルの推定を、各セグメントごとに試みようとしたが、セグメント 2において全く有意な値が得られなかった。この結果は、セグメント 2を構成しているサンプル全員の実際の利用交通手段が自家用車であり、セグメント 2を構成しているのは鉄道利用者がほとんどであることによるものであった。その R P での利用交通手段の内訳を表 3－4 に示す。今回行ったセグメンテーションは、定数項の符号の違いがセグメント分割の要因になっていると述べたが、定数項は導入した説明変数以外の要因の影響の新規高速船の利用意向の期待値であると

表 3－4 セグメント別 R P での利用手段

|         | セグメント 1 (負定数項) | セグメント 2 (正定数項) |
|---------|----------------|----------------|
| 自家用車利用者 | 9              | 4 2            |
| 鉄道利用者   | 3 1            | 0              |

考えられる。例えば新規高速船利用に伴う乗り換え数などの新規高速船の利用意向に対する期待値が定数項に現れているため、この結果は R P の利用交通手段と推定した個人モデルの定数項が相関関係にあることを示している。そこで、実際の利用交通手段別で定数項が異なっているかを検証するために、R P での利用交通手段別にデータをわけて、同じ説明変数を用いたモデルを推定した結果が表 3－5 である。この結果から明らかなように、セグメント 2と車利用者層とでは差があることが定数項の値の違いによって分かる。したがって、車利用者層が似たような嗜好を持っているとは断言できない。一方、セグメント 1と鉄道利用者では定数項に大きな差は見られず、鉄道利用者は全員がセグメント 1のメンバーであることを考えると、セグメント 1のパラメータは R P の利用交通手段に関わらず安定的であることを示している。つまりセグメント 1が新規高速船の利用意向に関して R P に依存せず同質的であるといえる。しかし、今回の S P 質問では実際の利用交通手段別に設定が異なり、新規高速船へのアクセス・イグレスが R P で選択した交通手段を使うという制約条件が課されているため、鉄道を使って新規高速船を利用する場合と車を使って利用する場合とでは、新規高速船の乗船場への経路の不明確さや乗り換えなどの利用意向への影響が、鉄道を使った場合の方がより大きく働き、その結果が定数項の同質性に結びついたと言うことが否定できない。つまり、実際の利用交通手段を用いるという制約条件が働いて、鉄道利用者全員の利用意向をある大きさだけ低下させてしまったという調査の設定条件による選好の変動が存在する可能性があるということである。一方、車利用者に対してはそのような制約条件による選好の変動の影響が小さく、純粋に利用意向の差異によって分割されていると考えることもできる。実際には鉄道利用者の新規高速船に対する利用意向が同質的なのか調査設計条件によるものなのかは、本研究の結果からでは断言できない。このことを明らかにするには、例えば今回のような制約（実際の利用交通手段を利用するという制約）を課した S P データと、制約を課さない S P データ（実際の利用交通手段



表 3－5 R P の選択手段別にサンプルをプールした S P 選択モデル

|             | Car            | Rail           |
|-------------|----------------|----------------|
| フェリー定数      | -0.0288 (-0.1) | -3.23 (-6.0)   |
| 総費用         | -0.488 (-5.8)  | -0.580 (-3.2)  |
| 運行頻度        | 1.43 (6.0)     | 3.59 (5.4)     |
| 総所要時間       | -0.0201 (-0.5) | -0.0839 (-1.3) |
| サンプル数       | 306            | 186            |
| $\bar{p}^2$ | 0.309          | 0.342          |

( ) 内は t 値

に関わらず同一条件の質問設定)を同時に収集して比較することによって可能となると考えられる。

### 3－4 個人の潜在的態度を考慮した行動モデル

#### 3－4－1 個人の態度と選択特性

態度とは前章でも述べたように消費者行動の意思決定過程において、意思決定者が代替案の属性や、代替案そのものに対して抱えている感情である。つまり、具体的には選好を表す効用関数のパラメータに影響しているものと考えられ、この異質性が選好パラメータの異質性と直接的に関連があると考えられる。通常、交通機関選択において所要時間や費用などの、どの要因を重視したかを尋ねることにより得られるデータは、このような交通行動の意思決定過程における潜在的な態度を、実際に観測可能な指標により定量的に測定したものであると考えることができる。本アプローチでは、このような潜在的態度変数が嗜好を表す効用パラメータに影響するという考えのもとに、観測可能な態度指標より個人の嗜好の違いを示すベクトルを求め、それに基づいたマーケットセグメンテーションを行うものである。本来ベネフィットセグメンテーションは、母集団に属する個人をそれぞれ独自のセグメントに帰属させることが、究極のセグメンテーションになると考えられる。しかし、3－2で述べたように、形成されるセグメントはマーケティング活動を行う意味のある大きさを持ったものにする必要があるため、嗜好が同質と思われる個人を同一のセグメントに帰属させ、母集団をたかだか数個のセグメントに

分割することが必要になる。そのための具体的手法として、本研究では、複数 S P データより推定した、個人ごとのパラメータベクトルに基づいたセグメンテーションを 3－3－3 で検討し、ここで、潜在的態度変数によるセグメンテーションを検討している。それぞれの方法は、図 1－1 に示された意思決定過程で嗜好の違いを表す態度の解明を異なるアプローチにより行うものである。つまり、前者は図 1－1 に示した態度を観測可能な選好意識から直接求めようとするアプローチであり、後者は態度を観測可能な態度指標から間接的に求めようとするものである。

#### 3－4－2 態度データの特性とその定性的分析

線形構造方程式モデルによる潜在的態度変数の推計に先だって、各個人がどのような要因を重視して交通機関選択を行ったかを把握するため、態度指標（交通機関選択理由）の単純集計を行う。アンケートでは以下の 9 項目について、今回の交通機関選択に関して重要視した項目を最大 3 つまで回答してもらう形式をとっている。3－4－1 でも述べたように 211 人から有効な回答が得られている。交通機関選択理由の質問に対する回答が容易であるため、個人パラメータによるアプローチのサンプル人数よりもかなり多いサンプル人数となった。

混雑が少ない (congestion)  
乗り換えが少ない (transfer)  
快適である (comfort)  
荷物を運ぶのが楽である (ease)  
静かである (silent)  
安全性が高い (safe)  
運賃・料金が安い (cheap)  
所要時間が短い (short)  
時間が正確である (relia)

集計結果を、図 3－4 ～図 3－6 に示す。図 3－4 は各項目を単純集計したものであり、約半数の人が所要時間の短さを重視していることがわかる。また、交通機関選択において所要時間と並んで特に重要視されられると思われる費用よりも混雑や快適性を重視していると回答した人が多いということが特徴として挙げられる。図 3－5 はどのような項目が同時に選択されてい



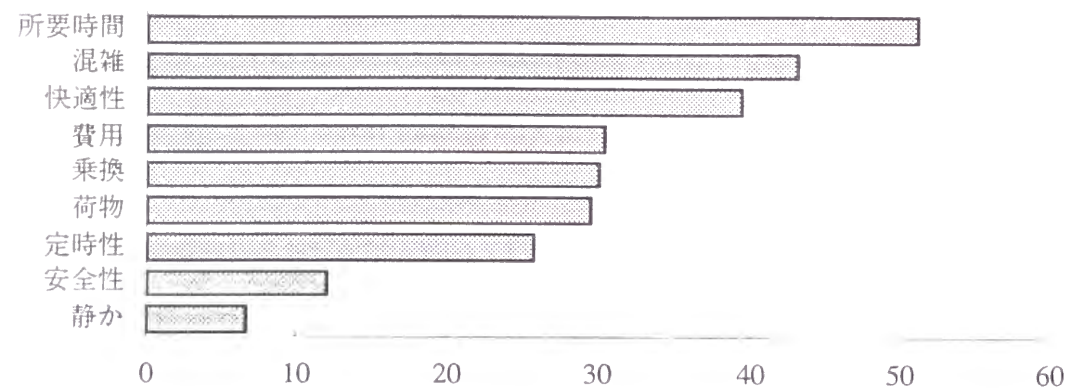


図 3-4 選択された態度指標の集計結果

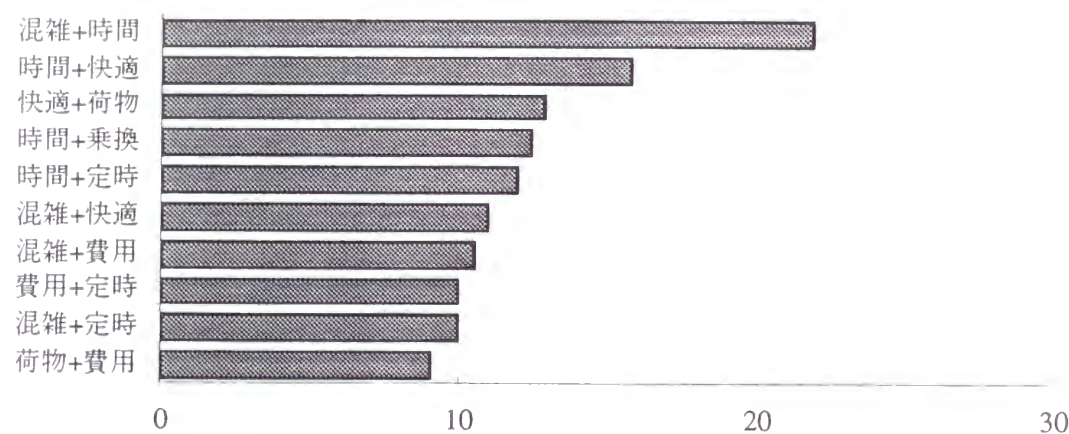


図 3-5 同時に選択された態度指標の組み合わせ上位 10 組

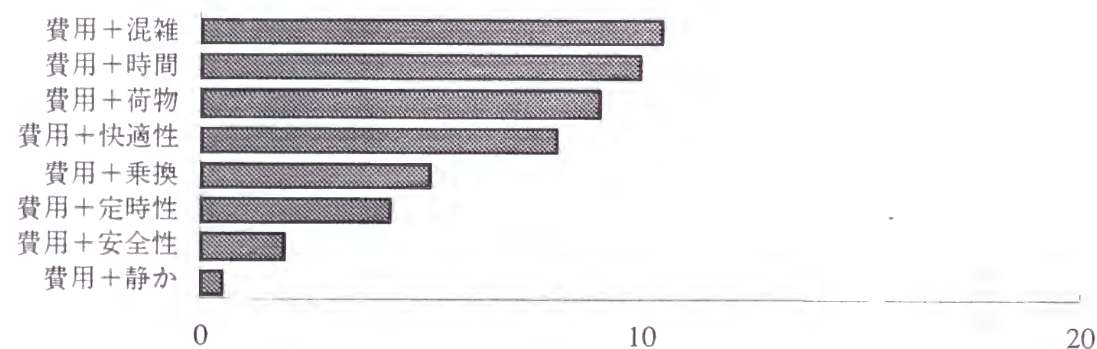


図 3-6 費用と同時に組み合わせて選択される態度指標

るのかを上位 10 位まで示したものである。当然の事ながら、所要時間の短さとペアになる項目が多いが、特に混雑と所要時間を同時に選択した人が多く、自家用車利用者は混雑の影響としては所要時間の増大を強く認識していることがうかがえる。所要時間を含まないものでは快適

性と荷物とを同時に選択している人が最も多い。また、費用は比較的均等に単純集計の上位の項目と同時に選択されている事が図 3-6 からわかる。

### 3-4-3 LISREL モデルによる態度データの分析

ここで、潜在的な態度を定量的な潜在的態度変数として推計する。その方法は、2 章で行った潜在的知覚値の導出と同様に、潜在的態度変数と観測可能な客観的変数（個人の社会経済属性および代替案の属性）との因果関係を表す構造方程式と、潜在的態度変数と観測可能な態度指標との関係を表す測定方程式とからなる線形構造方程式モデルを推定し、得られた未知パラメータの推定値から潜在的態度変数の推計値を計算する。以下に、線形構造方程式モデルの定式化および未知パラメータの推定方法、潜在的態度変数の推計値の計算方法について述べる。

構造方程式

$$\mathbf{w}^* = \mathbf{B}\mathbf{s} + \boldsymbol{\zeta} \quad (3.1)$$

ただし

$\mathbf{w}^*$ ：潜在的態度変数ベクトル

$\mathbf{s}$ ： $\mathbf{w}^*$ を形成する客観的説明変数ベクトル

$\mathbf{B}$ ：未知パラメータ行列

$\boldsymbol{\zeta} \sim \text{MVN}(0, \boldsymbol{\Psi})$ に従うランダム項（ $\boldsymbol{\Psi}$ は共分散行列）

測定方程式

$$\mathbf{Y} = \boldsymbol{\Lambda}\mathbf{w}^* + \boldsymbol{\varepsilon} \quad (3.2)$$

ただし

$\mathbf{Y}$ ：アンケートで得られた態度指標ベクトル

$\boldsymbol{\Lambda}$ ：未知パラメータ行列

$\boldsymbol{\varepsilon} \sim \text{MVN}(0, \boldsymbol{\Theta})$ に従うランダム項（ $\boldsymbol{\Theta}$ は共分散行列）

このモデルの未知パラメータの推定は、すべての変数が多変量正規分布しているという仮定のもとに、式 (3.1)、(3.2) から求められる観測可能な変数ベクトルの出現確率を尤度として、最尤推定法を用いて行う。このとき、潜在的態度指標がカテゴリカルデータである場合が多いため、PRELIS などの統計ソフトを用いて連続変数に変換して観測変数の共分散行列を求める。

推定された線形構造方程式モデルのパラメータを用いて、2章と同様に次式によって潜在的態度変数の推計値  $\bar{w}^*$  を計算する。

$$\bar{w}^* = Bs + \Psi \Lambda' (\Lambda \Psi \Lambda' + \Theta)^{-1} (Y - \Lambda Bs) \tag{3.3}$$

3-4-2で行った態度指標の集計分析より、本研究では潜在的態度として、所要時間重視態度と快適性重視態度（肉体的・精神的）を仮定することにする。ただし、双方とも費用を考慮した上でのものとする。そして、以下のように線形構造方程式モデルの特定化を行った。

構造方程式

所要時間重視態度

$$w_1^* = \beta_1 female + \beta_2 age20 + \beta_3 age50 + \beta_4 job + \beta_5 sightseeing + \beta_6 sports + \zeta_1 \tag{3.4}$$

快適性重視態度

$$w_2^* = \beta_7 female + \beta_8 age20 + \beta_9 age50 + \beta_{10} job + \beta_{11} sightseeing + \beta_{12} sports + \zeta_2 \tag{3.5}$$

ここに

|               |            |                    |            |
|---------------|------------|--------------------|------------|
| <i>female</i> | 1 ; 女性     | <i>job</i>         | 1 ; 主婦・無職  |
|               | 0 ; 男性     |                    | 0 ; それ以外   |
| <i>age20</i>  | 1 ; 20歳代以下 | <i>sightseeing</i> | 1 ; 観光目的   |
|               | 0 ; それ以外   |                    | 0 ; それ以外   |
| <i>age50</i>  | 1 ; 50歳代以下 | <i>sports</i>      | 1 ; スポーツ目的 |
|               | 0 ; それ以外   |                    | 0 ; それ以外   |

測定方程式

$$\begin{bmatrix} y_1(congestion) \\ y_2(transfer) \\ y_3(comfort) \\ y_4(ease) \\ y_5(silent) \\ y_6(cheap) \\ y_7(short) \\ y_8(relia) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_2 & \lambda_5 \\ 0 & 1 \\ 0 & \lambda_6 \\ 0 & \lambda_7 \\ \lambda_3 & \lambda_8 \\ 1 & 0 \\ \lambda_4 & 0 \end{bmatrix} \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \end{bmatrix} \tag{3.6}$$

表 3-6 LISRELモデルのパラメータ推定結果

$$\beta = \begin{bmatrix} (w_1^*) & (w_2^*) \\ -0.0326 & 0.0343 & (female) \\ (-0.7) & (0.8) \\ -0.104 & 0.0572 & (age20) \\ (-2.2) & (1.3) \\ -0.0313 & 0.0321 & (age50) \\ (-0.7) & (0.8) \\ -0.0071 & -0.0164 & (job) \\ (-0.2) & (-0.4) \\ 0.0223 & -0.648 & (sightseeing) \\ (0.5) & (-1.4) \\ 0.0915 & -0.0800 & (sports) \\ (1.9) & (-1.8) \end{bmatrix} \quad \Lambda = \begin{bmatrix} (w_1^*) & (w_2^*) \\ 0.854 & 0 & (congestion) \\ (4.1) & \\ 0.487 & 0.685 & (transfer) \\ (2.0) & (2.4) \\ 0 & 1 & (confort) \\ 0 & 1.13 & (ease) \\ & (4.1) \\ 0 & 0.541 & (silent) \\ & (2.7) \\ 0.459 & 1.00 & (cheap) \\ (1.6) & (3.2) \\ 1 & 0 & (short) \\ 0.962 & 0 & (relia) \\ (4.3) & \end{bmatrix}$$

推定値下の（）内はt値

上記のモデルを最尤推定した結果を表3-6に示す。構造方程式のパラメータ値のt値は低いものも多いが、測定方程式のパラメータ値はすべて正であり、所要時間重視態度についての費用が安いに対するパラメータ値を除く全てのパラメータ値が有意水準5%で有意であることから、2つの潜在的態度変数はそれぞれ費用を考慮した上での所要時間重視態度と快適性重視態度を表しているものと考えられる。所要時間重視態度は特に所要時間の短さや時間が正確であるといった態度指標へ大きく寄与しており、快適性重視態度は荷物運びの楽さや快適性、費用の安さといった態度指標に大きく寄与していることがパラメータ値から分かる。

### 3-4-4 態度変数によるセグメンテーション

SPデータは一種の実験データであるため操作性に優れ、意思決定における属性間のトレードオフ情報を把握しやすいため、本研究のように個人の異質性を捉えるという点においては優れていると言える。したがって、得られたセグメントの差異を捉えるためにSPデータを用い

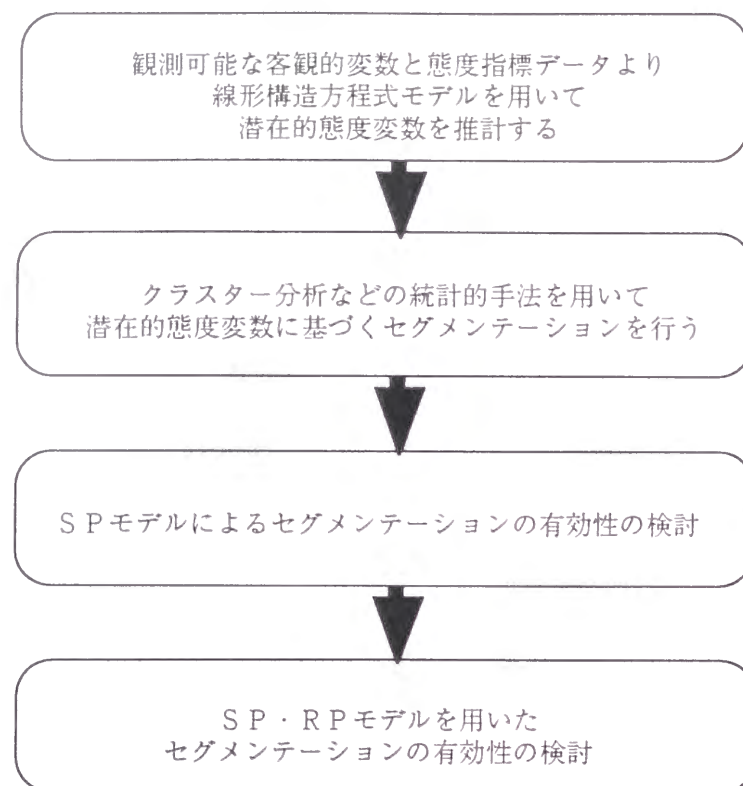


図3-7 潜在的態度変数によるセグメンテーションのフローチャート

ることでセグメンテーションの有効性に関する検討を行う。具体的には、各セグメントごとにデータをプールしてロジットモデルによりパラメータ推定を行い、セグメンテーションを行わないモデルとの比較を行う。潜在的態度変数に基づくアприオリセグメンテーションは、図3-7に示すフローにしたがって行われる。まず、得られた態度指標データと客観的変数（個人の社会経済属性およびトリップ属性）を用いて線形構造方程式モデルを推定し、(3.3)式を用いて潜在的態度変数の推計値（fitted value）を計算する。そして、クラスター分析によってセグメンテーションを行い、得られたセグメントごとに、各セグメントごとの嗜好の違いがより反映されと考えられるSPデータを用いてロジットモデルの推定を行う。その結果と、セグメンテーションを行わない場合の結果とを比較することによって、セグメンテーションの有効性の検討を行う。最後に、RP・SPモデルを推定しセグメンテーションの安定性の検討を行う。

線形構造方程式モデルのパラメータ推定値を用いて、式(3.3)より各個人についてそれぞれの潜在的態度変数の推計値を求め、散布図として表したのが図3-8である。見た目でもわか

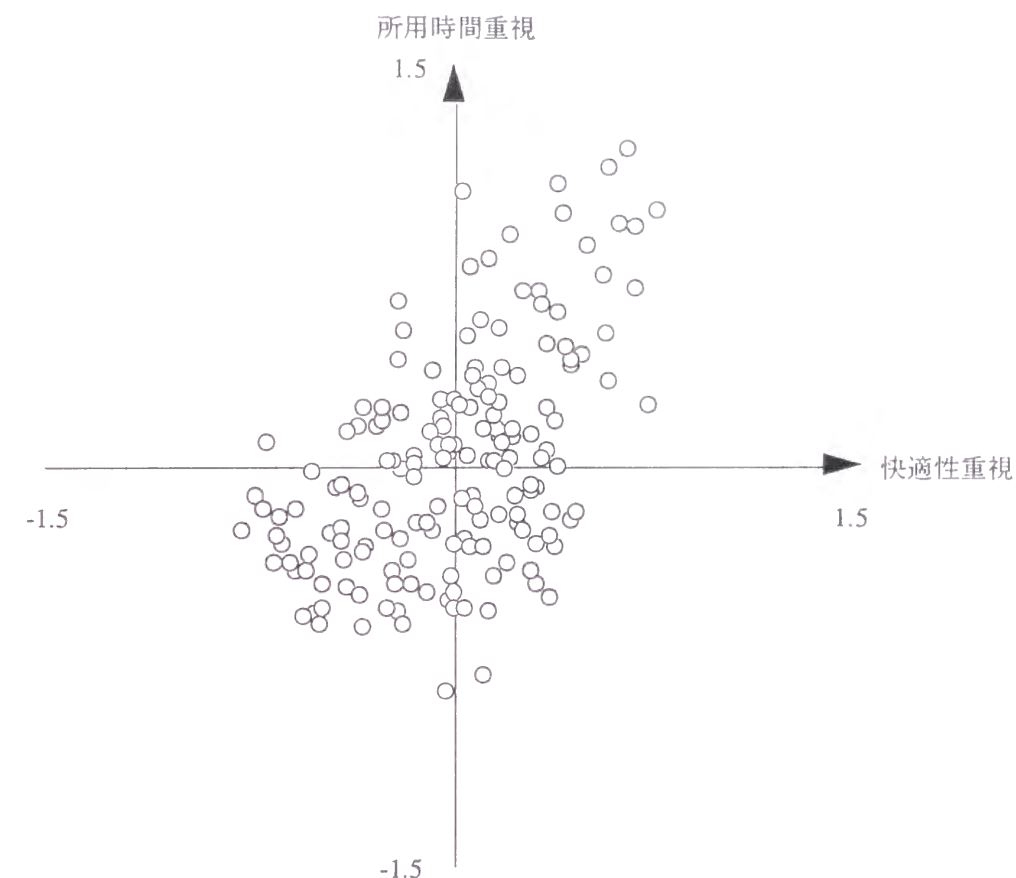


図3-8 潜在的態度変数の散布図

るように、この2つの潜在変数の間には正の相関が存在する。これは3-4-2で見たように、態度指標の間には、あるものが選ばれたときには特定の他の指標も選ばれていることが多いという、相関が存在することよるものと思われる。ここで、得られた潜在的態度変数をクラスター分析にかけたが、意味が明確なセグメントが抽出できなかった。ここで、潜在的態度変数本来の考えに立ち戻り、潜在的態度変数が個人の嗜好の違いを示し、効用パラメータに反映されるという考えに基づくならば、その大きさにセグメンテーションする、つまり、潜在的態度変数の比を基準にすることが妥当であるといえる。そこで本研究では、45度線を境界としてセグメンテーションを行った。その結果が図3-9である。45度線より上側が所要時間重視セグメントであり、下側が快適性重視セグメントとなる。

ここで、各セグメントの実際の利用交通手段の分担率を集計した結果が表3-7である。この表から、快適性重視セグメントでは全体よりも車の利用率が高く、所要時間重視セグメント



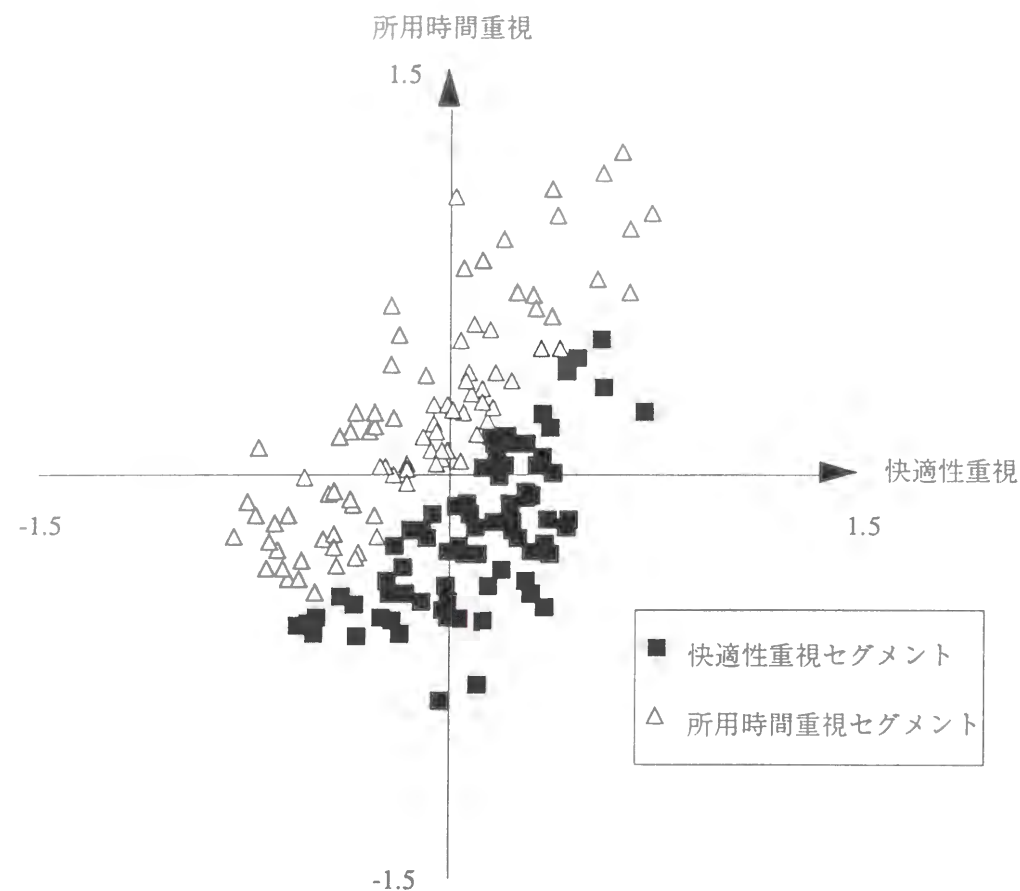


図3-9 潜在的態度変数比に基づいたセグメント

表3-7 態度変数に基づくセグメント内のRP利用交通手段分担率

|      | 全データ | 快適性重視セグメント | 所用時間重視セグメント |
|------|------|------------|-------------|
| 鉄道   | 36%  | 22%        | 49%         |
| 自家用車 | 64%  | 78%        | 51%         |

では全体よりも鉄道の利用率が高い傾向にあることが分かる。これは、快適性重視態度が荷物運びの楽しさや費用の安さといった車の長所となる態度指標に大きく寄与しており、所要時間重視態度が所要時間の短さや定時性といった鉄道の長所となる態度指標に大きく寄与していることから考えると妥当な結果であるといえよう。

図3-9より得られたセグメントごとにデータをプールして3-3-3と同様にロジットモデルを推定する。この際、個人パラメータアプローチで行った考察を考慮して、実際の利用交

通手段に関するダミー変数（1:鉄道利用者、0:車利用者）を加えて推定を行う。このダミー変数は、SPデータのバイアスである過去の行動の正当化を反映させるため（森川、Ben-Akiva（1992））だけではなく、前述した鉄道利用者固有の利用意向または実際の利用交通手段を利用するという制約によるSP選好の変動の大きさの測定を目的として導入する。パラメータ推定を行った結果が表3-8である。セグメンテーションを行わないモデルと比較して、 $\bar{\rho}^2$ 値はあまり向上していないが、尤度比検定により有意水準5%で各セグメント間のパラメータベクトルの等価性は棄却され、このセグメンテーションの有効性を示している。また、時間に対するパラメータ値が快適性重視セグメントにおいては有意でないのに対して、所要時間重視セグメントにおいては有意になっている。これは、交通機関の選択に所要時間を重視すると答えた人は、所要時間の変動に対して有意に反応することを示している。得られたパラメータの推定値から各セグメントの時間価値（総所要時間のパラメータ値/総費用のパラメータ値）を算定すると、快適性重視セグメントでは400円/時であり、所要時間重視セグメントでは1500円/時である。快適性重視セグメントの時間に対するパラメータ値が有意でないため、単純に比較することはできないものの、その差は大きい。また実際の利用交通手段に関するダミー変数に対するパラメータ値は予想どおり有意に負の値となっている。つまり、鉄道利用者は車利用者に

表3-8 態度変数に基づいたセグメント別のパラメータ推定結果

|                | 全データ                                      | 快適性重視<br>セグメント | 所用時間重視<br>セグメント |
|----------------|---|----------------|-----------------|
| フェリー定数項        | -0.684 (-2.3)                             | -0.469 (-1.2)  | -0.781 (-1.7)   |
| 総費用            | -0.426 (-8.7)                             | -0.380 (-6.7)  | -0.512 (-5.7)   |
| 運行頻度           | 1.39 (9.2)                                | 1.11 (5.9)     | 1.75 (7.4)      |
| 総所要時間          | -0.0753 (-3.1)                            | -0.0262 (-0.8) | -0.126 (-3.7)   |
| R P モードダミー     | -0.827 (-3.0)                             | -0.975 (-2.5)  | -1.02 (-2.4)    |
| 観測数            | 819                                       | 422            | 397             |
| $\rho^2$       | 0.282                                     | 0.300          | 0.284           |
| $\bar{\rho}^2$ | 0.273                                     | 0.275          |                 |
| $\chi^2$ 値     | $\chi^2 = 12.56 \chi_{0.05}^2(5) = 11.07$ |                |                 |

( ) 内はt値



比べその値だけ一律に新規高速船定数項が小さいということになる。快適性重視セグメントにおいても、所要時間重視セグメントにおいてもそのダミー変数に対するパラメータの絶対値は、新規高速船定数項の絶対値に比べ大きな値を取っており、実際の利用交通手段がどちらかであるかによってこの選択モデルに大きな影響を及ぼしていることが明らかにされたと言える。ここでもその影響が鉄道利用者に固有な新規高速船の利用意向なのか、SPデータの設計条件によるものなのかは明らかにできないが、ダミー変数を導入することによってその影響を考慮したセグメンテーションが行えることが示されたと言える。

続いて、3-3-3と同様にSPデータとRPデータを同時に用いたロジットモデルの推定を行う。用いた説明変数は、RPモデルに固有な変数として鉄道定数、SPモデルに固有な変数としてフェリー定数、運行頻度、表3-8と同様に実際の利用交通手段のダミー変数（RPダミー）である。また、RPモデルとSPモデルに共通な変数として総所要時間と総費用を用いてパラメータ推定を行った結果が表3-9である。ほとんどのパラメータ値が有意で、妥当な符号を持っていることが推定結果よりわかる。また、尤度比検定によって危険率5%でセグメント間のパラメータベクトルの等価性は棄却され、このセグメンテーションの有効性が確認できる。セグメント間での大きな違いは、スケールパラメータの値が所要時間重視セグメントでは1より大きい値であり、快適性重視セグメントでは1より小さい値となっていることである。すなわち、所要時間重視セグメントではSPデータに含まれるノイズがRPデータのそれよりも小さいということであり、より正確な属性間のトレードオフ情報がSPデータより得られていると考えられる。一方、快適性重視セグメントではその逆である。また、パラメータの推定値から得られる時間価値はそれぞれ所要時間重視セグメントでは2100円/時、快適性重視セグメントでは1300円/時であり、その差に大きな違いがみられる。しかし、このようなセグメントごとのパラメータの相対的な値は、SPデータだけを用いたモデルからそれほど変化がないことから、SP・RPデータを同時に用いた場合でも態度変数に基づいたセグメントは異種データの混合に対して安定していると言える。以上の結果から、実際の行動結果のデータを統合したモデルにおいても、態度変数に基づいて分けられたセグメント間で効用パラメータの違いが有意に存在しており、このセグメンテーションの安定性が示されたといえる。特に実際の利用交通手段のダミー変数を導入することによって、鉄道利用者に固有な新規高速船への利用意

表3-9 潜在的態度変数に基づいたセグメント別のSP・RPモデルの推定結果

|                | 全データ           | 快適性重視<br>セグメント   | 所要時間重視<br>セグメント            |
|----------------|----------------|------------------|----------------------------|
| フェリー定数 (SP)    | -0.697 (-2.2)  | -0.918 (-1.6)    | -0.643 (-1.5)              |
| 鉄道定数 (RP)      | -0.744 (-4.9)  | -1.50 (-5.3)     | -0.198 (-0.9)              |
| 総費用            | -0.281 (-5.6)  | -0.394 (-3.5)    | -0.253 (-3.5)              |
| 運行頻度 (SP)      | 1.00 (3.9)     | 1.26 (2.7)       | 0.984 (2.5)                |
| 総所要時間          | -0.0766 (-3.1) | -0.0822 (-1.9)   | -0.0873 (-2.2)             |
| RPダミー (SP)     | -0.436 (-2.6)  | -0.816 (-1.9)    | -0.400 (-2.0)              |
| スケールパラメータ      | 1.40 (4.1)     | 0.894 (2.9)      | 1.80 (2.6)                 |
| サンプル数          | 1030           | 524              | 506                        |
| $\rho^2$       | 0.257          | 0.309            | 0.244                      |
| $\bar{\rho}^2$ | 0.247          |                  | 0.258                      |
| $\chi^2$ 値     |                | $\chi^2 = 29.63$ | $\chi_{0.05}^2(7) = 14.07$ |

( ) 内はt値

向なのかSPデータの設計条件による選好の変動かは断定できないが、その影響をも考慮した有効でSP・PR両データに対して安定したセグメントを形成することが示された。

### 3-4-5 潜在クラス分析と潜在セグメント分析

個人の効用パラメータの異質性を考えた場合、根源的には母集団に属するサンプルの数と同じだけセグメントが存在し、その個人（セグメント）ごとに効用関数の形が異なっていると考えられる。このように異なった効用パラメータを持つ母集団に対し、セグメント内で確定的に同一パラメータを与えるアプローチは、その母集団内のパラメータの分布がきれいな単峰型である場合に限り良い近似を与えるが、そうでない場合はかえってセグメントの特性をわかりづらくする原因となりかねない。マーケティング分野での研究報告によると、個人の効用パラメータの分布は決して分割可能な多峰型とは限らないというものもある。そこで、3-4-3で求めた態度変数を用い、個人のパラメータを確率的に推定する潜在クラス分析を行う。潜在クラス分析とは個人の異質性を考慮するモデルとして近年マーケティング分野で研究されてい

るモデルで、多変量解析の分野でカテゴリカルデータの因子分析手法として開発されたものにベースを置いており、マーケティングの分野でマーケット構造の分析に用いられるようになってきた。この初期の適用研究としては Clogg (1981)、Clogg and Sawyer (1981)、Andersen (1982)、Bergan (1983) などがあげられる。潜在クラスモデルは効用パラメータの異質性だけでなく様々な局面で用いられている。マーケットのクラスター分析の一手法つまり、同質な意思決定者分類方法の一つとして用いられることもあれば、先に述べた因子分析手法的な潜在変数を仮定するモデルとしての扱い、つまり、マーケット構造を規定するカテゴリー変数の背後の潜在変数を規定する手法として用いられることもある。ここで、柳井ら (1990)、Dillon and Kumar (1994) を参考に一般的なモデルを簡単に説明する。2 値データ  $x$  の反応傾向のばらつきを説明するため、いくつかの潜在的なクラス (集団) を想定する。クラスの数  $c$  とし、それぞれのクラスを  $k$  で表す。そのときある反応パターン  $x$  を得る確率は

$$P(\tilde{x} = x) = \sum_{k=1}^c P(x | k) P(k) \quad (3.4)$$

によって表される。これは簡単に多値データにも適用することができる。このモデルでは、あるクラスに属する確率  $P(k)$  の与え方は任意であり、様々なものが提案されているが、Formann (1982) は次式に示すように、ロジスティックモデルを提案している。

$$P(k) = \frac{\exp(w_k)}{\sum_{k=1}^K \exp(w_k)} \quad (3.5)$$

ここで、セグメント帰属確率、選択確率をいずれもロジットモデルによって表現した潜在セグメント分析手法の定式化を示す。

個人  $n$  がセグメント  $s$ ,  $s=1,2,\dots,S$  に帰属しているときの効用関数を

$$U_{n|s}^* = \beta_s X_{ni} + v_{n|s} \quad (3.6)$$

とおく。ただし、

$U_{n|s}^*$ : セグメント  $s$  に帰属した場合の個人  $n$  の選択肢  $i$  に対する効用

$\beta_s$ : セグメント  $s$  の効用パラメータベクトル

$X_{ni}$ : 個人  $n$  の属性と選択肢  $i$  の属性のベクトル

$v_{n|s}$ : セグメント  $s$  に属する個人  $n$  の選択肢  $i$  に対する効用の誤差項

このモデルは、各潜在セグメントごとに異なるパラメータを与えることによって、母集団内の効用パラメータの異質性を表している。ここで  $v_{n|s}$  が独立で同一のガンベル分布に従うと仮定すると、セグメント  $s$  が与えられた条件付きの選択確率は

$$P_{n|s} = \frac{e^{\beta_s X_{ni}}}{\sum_{j \in C_n} e^{\beta_s X_{nj}}} \quad (3.7)$$

と表される。

ただし、 $C_n$  は個人  $n$  の選択肢集合

母集団中の潜在セグメントの数が既知であるとするとき、各個人が潜在的なセグメントに属する確率を決定する変数として  $Y^*$  を仮定する。 $Y^*$  は各個人の態度指標などにより推定された潜在的態度変数と各個人の属性の線形和で与えられる。

$$Y_{ns}^* = \Gamma_{Gs} G_{ns}^* + \Gamma_{zs} z_{ns} + \delta_{ns} \quad (3.8)$$

$Y_{ns}^*$ : 個人  $n$  のセグメント  $s$  に対する帰属度関数

$G_{ns}^*$ : 個人  $n$  の潜在セグメント  $s$  の帰属度に影響を与える態度変数

$z_n$ : 個人  $n$  の潜在セグメント  $s$  の帰属度に影響を与える社会経済属性ベクトル

$\Gamma_{zs}$ ,  $\Gamma_{Gs}$ : 未知パラメータベクトル

$\delta_{ns}$ : 個人  $n$  のセグメント  $s$  に対する帰属度変数の誤差項

このとき Luce の選択公理 (Luce (1959)) に従い、セグメント間の帰属比率が2つのセグメント間の帰属度の大小関係によって決定されると仮定する、つまり  $\delta_{ns}$  が独立で同一のガンベル分布に従うと仮定すると、各セグメントの帰属確率は次式で表される。

$$W_{ns} = \frac{e^{\Gamma_{zs} z_n}}{\sum_{k=1}^S e^{\Gamma_{ks} z_n}} \quad (3.9)$$

意思決定者  $n$  が潜在セグメント  $s$  に帰属し、代替案  $i$  を選択する同時確率は、局所独立性の仮定により以下のような形で表すことが可能になる。

$$P_{nis} = P_{n|s} W_{ns} \quad (3.10)$$

この条件付き確率の周辺和を取ることで、意思決定者  $n$  が代替案  $i$  を選択する確率が得られる。

$$P_{ni} = \sum_{s=1}^S P_{ni|s} W_{ns} \quad (3.11)$$

ただし、 $S$  は潜在セグメントの総数

Swait (1993) はこのロジット型モデルを、化粧品ブランドのブランドチョイスデータに適用している。ここでは、メンバーシップ尤度関数  $Y_{ns}^*$  を以下のように規定している。

$$Y_{ns}^* = \tilde{\Gamma}_{ps} G_{np}^* + \tilde{\Gamma}_{as} G_{na}^* + \tilde{\Gamma}_{zs} X_{ns} + \zeta_{ns} \quad (3.12)$$

$$\tilde{X}_{np} = B_p G_{np}^* + \zeta_{np} \quad (3.13)$$

$$\tilde{X}_{na} = B_a G_{na}^* + \zeta_{na} \quad (3.14)$$

$Y_{ns}^*$  : 個人  $n$  とセグメント  $s$  に関するメンバーシップ尤度関数

$\tilde{X}_{np}$ 、 $\tilde{X}_{na}$  : 観測される個人の知覚値指標と態度指標のベクトル

$G_{np}^*$  : 個人の潜在知覚値のベクトル

$G_{na}^*$  : 個人の潜在態度のベクトル

$X_{ns}$  : 観測される個人の社会人口統計学特性のベクトル

$\tilde{\Gamma}_{ps}, \tilde{\Gamma}_{as}, \tilde{\Gamma}_{zs}, B_p, B_a$  : 推定されるパラメータベクトル

$\zeta_{ns}, \zeta_{np}, \zeta_{na}$  : ランダム項

このとき、個人レベルでのセグメントメンバーシップ関数は以下になる。

$$Y_{ns}^* = \gamma_s Z_n + \zeta_{ns}, \quad s=1, \dots, S, \quad (3.15)$$

$$Z_n = \begin{bmatrix} \tilde{X}_{np} \\ \tilde{X}_{na} \\ X_n \end{bmatrix} \quad (3.16)$$

$$\gamma_s = (\tilde{\Gamma}_{ps}, \tilde{\Gamma}_{as}, \tilde{\Gamma}_{zs}) \quad (3.17)$$

ランダム項  $\zeta_{ns}, \zeta_{np}, \zeta_{na}$  が相互に独立で同一の分散をもったガンベル分布（スケールパラメータ  $\alpha$ ）であるとする個人  $n$  がセグメント  $s$  に属する確率は式 (3.18) より得られる。

$$W_{ns} = \frac{e^{\alpha \gamma_s Z_n}}{\sum_{k=1}^S e^{\alpha \gamma_k Z_n}} \quad (3.18)$$

$$V(\zeta_{ns}) = \pi^2 / 6 \alpha^2 \quad (3.19)$$

潜在的なクラス（セグメント）を考慮したモデルとしてマーケティング分野で適用されている片平（1991）の LOGMAP-J モデルがある。これは選好データもしくは選択データだけからブランドのマップと回答者個々のポイントを同時に定めるジョイント・スペース分析のモデルであり、ジョイント・スペース分析は「選好の内的分析」ともいわれていると同時に、各ブランドのマーケティング諸変数（値引き、マス広告等）に対する各個人の反応を測定しようとする市場反応分析モデルでもある。具体的には、潜在的なセグメントが与えられたとき選択肢  $i$  が選択される確率を多項ロジットモデルで表し、それぞれに規定したモデルより得られるセグメント帰属確率との同時確率を求めることで選択肢  $i$  の選択確率を推定している。第  $s$  セグメントの効用関数は次のように規定される。

$$U_{sjt} = \alpha_s^* x_j^* + \alpha_s^{**} x_{jt}^{**} + \varepsilon_{sjt} \quad (3.20)$$

$x_j^*$  : ブランド  $j$  のマップ上の座標ベクトル

$x_{jt}^{**}$  : ブランド  $j$  のマップ上のマーケティング変数ベクトル

$\varepsilon_{sjt}$  : ランダム項

$\alpha_s^*, \alpha_s^{**}$  : パラメータベクトル

ある特定の個人  $k$  は  $S$  個のセグメントに確率的に所属するものとし、個人  $k$  がセグメント  $s$  に所属する確率を  $w_{ks}$  とする。 $\alpha_s = (\alpha_s^*, \alpha_s^{**})$  とすると個人  $k$  がパラメータ  $\alpha_s (s=1, \dots, S)$  を持つ  $S$  個のセグメントに属する所属確率  $w_{ks}$  は次のようになる。

$$w_{ks} = L(\alpha_s | y_k) / \sum_{u=1}^S L(\alpha_u | y_k) \quad (3.21)$$

$$L(\alpha_s | y_k) = \prod_j \prod_t P_{sjt}(\alpha_s)^{y_{kjt}} \quad (3.22)$$

$$P_{sjt} = \exp V_{sjt} / \sum_h \exp V_{shj} \quad (3.23)$$

$P_{sjt}$  :  $t$  期においてセグメント  $i$  の個人がブランド  $j$  を選択する確率

$L(\alpha_s | y_k)$  : 個人  $k$  のブランド選択データ  $y_k$  が与えられたもとのパラメータ  $\alpha_s$  の尤度



本研究では、ロジット型関数モデルによって特定のクラスへ属する確率を規定し、効用パラメータに潜在的クラスを仮定するモデルを用いて、意思決定者の異質性を考慮したモデルを構築する。ここで、このモデルの本質がセグメント分析を確率的に行うことと、潜在的クラスという名称が従来土木計画学の分野であり一般的でないことから、特に潜在セグメントモデルと定義し、3-3-3までで、2つのセグメンテーションアプローチの適用を行った新規高速船に関する複数SPデータに適用し、その有効性を探る。

#### 3-4-6 態度変数を用いた潜在セグメント分析

3-4-3で求めた態度変数を用いて、潜在セグメント分析と離散型選択モデルを組み合わせたモデルの推定を行う。ここで、同一個人から繰り返し採取するデータを用いて離散型選択モデルを推定する場合、Omitted Variableなどの影響により効用関数の確率項には、各選択肢ごとに意思決定者個人の系統的な特徴が含まれ、同一個人から複数のデータを得てモデルを推定した場合、誤差項間に系統的な相関が生じることがある。そこでこの相関を解消するために、Morikawa (1994)が提案する効用関数の誤差項を個人に共通の $\lambda$ と真にランダムなホワイトノイズを表す項 $v$ に分解し、時点間での誤差項の相関を解消するモデルを導入する。この系列相関モデルについては次章で詳しく解説を行う。

帰属するセグメント $s, s=1, \dots, S$ が与えられたとき、ランダム効用理論に基づく効用関数は

$$U_{nlis} = \alpha_s X_{nli} + \lambda_{ni} + v_{nlis} \quad (3.24)$$

ただし、

$U_{nlis}$  : セグメント $s$ に帰属した個人 $n$ の $l$ 回目のSP質問の選択肢 $i$ に対する効用

$\alpha_s$  : セグメント $s$ の効用パラメータベクトル

$X_{nli}$  :  $l$ 回目のSP質問での個人 $n$ の属性と選択肢 $i$ の属性のベクトル

$\lambda_{ni}$  : SP質問の回数やセグメントにかかわらず、同一個人で共通の値を取る誤差を表す系列相関項

$v_{nlis}$  : 効用の誤差項

のようになる。また各セグメントへの帰属確率は個人の態度変数を $\hat{w}_{ns}^*$ として帰属度関数を

$$G_{ns}^* = \Gamma_s \hat{w}_{ns}^* + \delta_s \quad (3.25)$$

とおくと次式で表される。

$$H_{ns} = \frac{e^{\Gamma_s \hat{w}_{ns}^*}}{\sum_{k=1}^S e^{\Gamma_s \hat{w}_{nk}^*}} \quad (3.26)$$

このとき $\lambda$ が与えられたとき個人 $n$ が $l$ 回目のSP質問で選択肢 $i$ を選ぶ条件付き確率は、潜在セグメントモデルを用いると(3.11)式に従って

$$P_{ni}(i_l | \lambda_{ni}) = \sum_{s=1}^S P_{nls}(i_l | \lambda_{ni}) H_{ns} \quad (3.27)$$

となる。

ここで、二項選択の場合、ある個人 $n$ が $l$ 回目のSP質問でそれぞれ選択肢 $i$ を選択する同時確率は、系列相関項 $\lambda_n$ の分布形を与え $\lambda_n$ に対する数学的期待値をとることにより得られる。

$$P_n(i_1, i_2, \dots, i_L) = \int_{-\infty}^{+\infty} P_{n1}(i_1 | \lambda_n) \cdot P_{n2}(i_2 | \lambda_n) \cdots P_{nL}(i_L | \lambda_n) f(\lambda_n) d\lambda_n \quad (3.28)$$

$f(\lambda_n)$  :  $\lambda_n$ の同時確率密度関数

これを個人の尤度関数として、最尤推定法により未知パラメータを推定する。この2段階推定は一致性のある推定量を与える(Amemiya (1985))。

3-4-3までに示したアプリオリセグメント手法はいずれも同一セグメント内では効用パラメータが同一という仮定が存在したが、このモデルの特徴として、セグメント帰属度変数に応じてセグメント帰属確率が変動するため、各個人の効用パラメータは、推定値に現れたセグメントごとの推定値を、各帰属度をパラメータとして内分する点になっている。言い換えれば、各個人ごとの効用パラメータを帰属度ベクトルを構成する変数によって構造化していることになる。このため推定に要する労力を低減して、各個人ごとのパラメータを推定するモデルになっていると言える。

ここで、3-4-3で求めた潜在的態度変数をセグメント帰属変数ベクトル $\hat{w}^*$ として、潜在セグメントモデルを推定した結果が表3-10に示してある。3-3-3で行った個人モデルを用いたセグメンテーション分析で、実際に利用した交通機関とSPでの選好に強い相関が存在する可能性が示されたため、その影響をモデル化するために効用関数にRPモードダミーを



表3-10 SPデータを用いた潜在セグメントモデルの推定結果

| 変数名            | 制約モデル          |  | 潜在セグメントモデル     |                       |
|----------------|----------------|--|----------------|-----------------------|
|                | 全データ           |  | セグメント1         | セグメント2                |
| フェリー定数         | -1.98 (-9.7)   |  | -2.63 (-7.2)   | 7.46 (4.7)            |
| 総費用            | -0.145 (-7.0)  |  | -0.0840 (-2.1) | -3.01 (-6.0)          |
| 運行頻度           | 1.40 (10.4)    |  | 1.96 (7.9)     | 1.08 (2.6)            |
| 総所要時間          | -0.0940 (-4.3) |  | -0.0671 (-1.6) | -0.779 (-5.0)         |
| 所用時間重視         |                |  | -0.968 (-1.4)  |                       |
| 快適性重視          |                |  | -0.147 (-0.3)  |                       |
| 観測数            | 819            |  | 819            |                       |
| $\bar{\rho}^2$ | 0.202          |  | 0.279          |                       |
| $\chi^2$ 値     |                |  | $\chi^2=265$   | $\chi^2_{0.05}=20.28$ |

() 内はt値

導入した。表3-10中で示した「各パラメータが母集団中で一定であるという制約を課したモデル」を使って潜在セグメントモデルの効用関数の各セグメントのパラメータが同一であるという制約を帰無仮説としたカイ2乗検定を行った結果、その帰無仮説は危険率5%で棄却された。表中に示された帰属度関数のパラメータの推定結果は、各セグメントへの帰属確率が快適性重視態度が大きくなるにつれてセグメント1への帰属確率が高まり、所用時間重視態度が大きくなるとセグメント2への帰属確率が高まることを示している。つまり、図3-9で45度線より上に離れるほど個人のパラメータはセグメント2の値に近づき、45度線より下に離れるとセグメント1のパラメータに近づく。45度線付近のサンプルはセグメント1と2のちょうど中間値をとっている。つまり、快適性重視がより大きい人は、運行頻度と総費用が有為に働き、フェリー定数が負である人で、所要時間重視態度の大きい人は、総所要時間、運行頻度、総費用いずれも有意でフェリー定数が正である人である。

またRPモードダミーを入れて推定した結果から(表3-11)セグメント1は公共交通利用者が新規高速船に対し利用意向の低いセグメントであることが読みとれるが、それほど推定値に変化は見られなかった。

表3-11 SPデータを用いた潜在セグメントモデルの推定結果(RPモードダミー)

| 変数名            | 制約モデル          |  | 潜在セグメントモデル     |                       |
|----------------|----------------|--|----------------|-----------------------|
|                | 全データ           |  | セグメント1         | セグメント2                |
| フェリー定数         | -1.86 (-7.7)   |  | -1.96 (-4.2)   | 7.62 (4.6)            |
| 総費用            | -0.162 (-6.1)  |  | -0.160 (-3.4)  | -3.09 (-5.8)          |
| 運行頻度           | 1.39 (9.6)     |  | 1.96 (8.1)     | 0.949 (2.2)           |
| 総所要時間          | -0.0989 (-4.6) |  | -0.0729 (-1.6) | -0.819 (-4.2)         |
| RPモードダミー       | -0.231 (-1.1)  |  | -1.10 (2.3)    | -0.993 (-1.0)         |
| 所用時間重視         |                |  | -1.04 (-1.6)   |                       |
| 快適性重視          |                |  | -0.0647 (-0.1) |                       |
| 観測数            | 819            |  | 819            |                       |
| $\bar{\rho}^2$ | 0.212          |  | 0.282          |                       |
| $\chi^2$ 値     |                |  | $\chi^2=265$   | $\chi^2_{0.05}=20.28$ |

() 内はt値

続いて、この分析手法を拡張してSP・RP両方のデータに適用した。複数SPデータに限らず、SP、RPも同一個人から繰り返して採取したデータになるので、その誤差項には系列相関が生じる可能性がある。そこで系列相関を考慮した定式化を行う。先に複数SPデータに対して定式化された(3.24)式から(3.28)式は以下のように変形される。

$$U_{nli}^{SP} = \alpha_s X_{nli}^{SP} + \beta_s Z_{nli} + \lambda_{ni} + v_{nli}^{SP} \quad (3.29)$$

$$U_{nli}^{RP} = \alpha_s X_{ni}^{RP} + \lambda_{ni} + v_{nli}^{RP} \quad (3.30)$$

ただし、

 $U_{nli}^{SP}$  : セグメントsに帰属した個人nのl回目のSP質問の選択肢iに対する効用

 $U_{nli}^{RP}$  : セグメントsに帰属した個人nのRP質問の選択肢iに対する効用

 $X_{nli}^{SP}$  : l回目のSP質問での個人nの選択肢iのRPにも存在する属性のベクトル

 $Z_{nli}$  : l回目のSP質問での個人nの選択肢iのSP特有の属性のベクトル

 $X_{ni}^{RP}$  : RP質問での個人nの選択肢iの属性のベクトル

 $\alpha_s, \beta_s$  : セグメントsの効用パラメータベクトル

 $\lambda_{ni}$  : SP・RP質問、セグメントにかかわらず、同一個人で共通の値を取る誤差を表す系

列相関項

$v_{nlls}^{SP}$  : S P 質問の効用の誤差項

$v_{nlls}^{RP}$  : R P 質問の効用の誤差項

よく知られているように、ロジットモデルを用いて推定されるパラメータは、ガンベル分布の分散を定めるスケールパラメータとの積の形で得られる (Ben-Akiva and Lerman (1986))。そのため、異なる状況のもとで得られたデータを用いて、ロジットモデルを推定した場合には、誤差が同一の分散であるという仮定が存在することになる。この仮定が保証されない場合には、どちらかの誤差項を基準とした分散のスケールを基準化する必要がある。そこで、(3.31) 式のようにスケールパラメータ  $\mu$  を導入して、2つの誤差項のスケールを調整する。ここでは、以下のような関係が成立している。

$$\text{var}(v_{nlls}^{SP}) = \mu^2 \text{var}(v_{nlls}^{RP}) \quad (3.31)$$

また各セグメントへの帰属確率は帰属度関数を S P モデルと同様に

$$G_{ns}^* = \Gamma_s \hat{w}_{ns}^* + \delta_s \quad (3.25)$$

とおき、ランダム効用理論と同様の仮定と誤差項に独立で同一のガンベル分布を仮定することで次式で表される。

$$H_{ns} = \frac{e^{\Gamma_s \hat{w}_{ns}^*}}{\sum_{k=1}^S e^{\Gamma_s \hat{w}_{nk}^*}} \quad (3.26)$$

このとき個人  $n$  が  $l$  回目の S P 質問で選択肢  $i$  を選ぶ条件付き確率は

$$P_{nl}^{SP}(i | \lambda_n) = \sum_{s=1}^S P_{nlls}^{SP}(i | \lambda_n) H_{ns} \quad (3.32)$$

となる。同様に R P 質問で選択肢  $j$  を選ぶ確率は

$$P_n^{RP}(j | \lambda_n) = \sum_{s=1}^S P_{nls}^{RP}(j | \lambda_n) H_{ns} \quad (3.33)$$

となる。ここで、ある個人  $n$  が  $l$  回目の S P 質問でそれぞれ  $i$  を選択し、R P で選択肢  $j$  を選択する同時確率は、系列相関項  $\lambda_n$  の分布形を与え、 $\lambda_n$  に対する数学的期待値をとることにより得られる。

$$P_n(i_1, i_2, \dots, i_L, j) = \int_{-\infty}^{+\infty} P_{n1}^{SP}(i_1 | \lambda_n) \cdot P_{n2}^{SP}(i_2 | \lambda_n) \cdots P_{nL}^{SP}(i_L | \lambda_n) P_n^{RP}(j | \lambda_n) \cdot f(\lambda_n) d\lambda_n \quad (3.28)$$

$f(\lambda_n)$  : の同時確率密度関数

これを個人の尤度関数として、最尤推定法により未知パラメータを推定する。この2段階推定もまた一致性のある推定量を与える。

この定式化に基づいて潜在セグメントモデルを推定した結果が下の表に示してある。S P モデルと異なり、R P データを同時に用いているため、R P モードダミーは特定化の段階ではずしたものと、効用関数に R P モードダミーを導入して推定した結果をそれぞれ表 3-12、表 3-13 に示した。それぞれ潜在セグメントを仮定しないモデルの推定結果を比較のため同一の表内に載せてある。両者の推定値より潜在セグメントの考慮にかかわらず、R P モードダミーの導入は他のパラメータの推定値にそれほど影響を与えていないことがうかがえる。ただし、潜在セグメント 1 においては R P モードダミーは有意となっているため、適合度が少し向上したが、R P モードダミーをこの S P・R P モデルに導入する必要性は薄いと考えられる。表 3-12、表 3-13 それぞれに示した各パラメータが母集団中で一定であるという帰無仮説に

表 3-12 態度変数を用いた S P・R P データの潜在セグメント分析 (R P モードダミー無し)

| 変数名             | 制約モデル         | 潜在セグメントモデル    |                       |
|-----------------|---------------|---------------|-----------------------|
|                 | 全データ          | セグメント 1       | セグメント 2               |
| 定数項             | 0.177 (1.44)  | 0.157 (1.7)   | 4.38 (5.2)            |
| 費用              | -0.215 (-6.8) | 0.0919 (2.7)  | -2.11 (7.3)           |
| 所要時間            | 0.103 (4.6)   | 0.150 (4.6)   | -0.230 (-3.1)         |
| 頻度              | 0.986 (7.5)   | 2.7 (7.2)     | 1.96 (6.6)            |
| スケールパラメータ $\mu$ | 1.61 (7.8)    | 3.84 (7.3)    |                       |
| 所用時間重視          |               | -0.566 (-1.0) |                       |
| 快適性重視           |               | -1.14 (-2.1)  |                       |
| 観測数             | 1030          | 1030          |                       |
| $\bar{\rho}^2$  | 0.156         | 0.280         |                       |
| $\chi^2$ 値      |               | $\chi^2=189$  | $\chi_{0.05}^2=20.28$ |

( ) 内は t 値

表 3-13 態度変数を用いた S P・R P データの潜在セグメント分析 (R P モードダミー有り)

| 変数名             | 制約モデル          | 潜在セグメントモデル    |                       |
|-----------------|----------------|---------------|-----------------------|
|                 | 全データ           | セグメント 1       | セグメント 2               |
| 定数項             | 0.173 (1.4)    | 0.149 (1.5)   | 4.14 (4.8)            |
| 費用              | -0.218 (-5.6)  | 0.0941 (3.1)  | -2.01 (-6.5)          |
| 所要時間            | 0.101 (4.2)    | 0.145 (5.0)   | -0.221 (-3.0)         |
| 頻度              | 0.990 (7.5)    | 3.27 (6.7)    | 1.88 (6.4)            |
| R P モードダミー      | -0.0429 (-0.1) | -1.48 (-2.3)  | 0.389 (0.3)           |
| スケールパラメータ $\mu$ | 1.59 (6.6)     | 3.80 (6.8)    |                       |
| 所用時間重視          |                | -0.604 (-1.1) |                       |
| 快適性重視           |                | -1.17 (-2.3)  |                       |
| 観測数             | 1030           | 1030          |                       |
| $\bar{\rho}^2$  | 0.156          | 0.291         |                       |
| $\chi^2$ 値      |                | $\chi^2=193$  | $\chi^2_{0.05}=20.28$ |

( ) 内は t 値

対するカイ 2 乗検定により、効用関数のパラメータが同一であるという制約条件の妥当性は危険率 5 % で棄却された。表中に示された帰属度関数のパラメータの推定結果は、S P モデルと同様に所要時間重視と快適性重視の差が正の方向に大きくなるにつれてセグメント 2 への帰属確率が高まり、逆に負の方向へ大きくなるとセグメント 1 への帰属確率が高まることを示している。それぞれのセグメントは、制約モデルと比較して、所要時間と費用のパラメータが、負と正に分かれ、定数項や頻度のパラメータにも違いが見られた。これらの一連のパラメータの傾向は、純粋に効用パラメータの異質性によるものだけでなく、非観測異質性による面もあると考えられる。

### 3-5 本章のまとめ

本研究では、個人の異質性を非集計離散型選択モデルに導入する簡便な方法としてのセグメンテーションを、個人の嗜好の違いに基づいて行うための手法を 2 つ提示し、事例研究により

その有効性の分析を行った。それに引き続いて、同様に潜在クラス分析を応用した潜在セグメントモデルを提案し、同じデータを用いて事例研究を行った。その結果得られたそれぞれのモデルの成果および今後の課題は以下にまとめられる。

個人パラメータに基づくセグメントを、S P データに適用することによって、個人の異質性を反映したセグメントを得ることができ、モデルの適合度が向上することを示した。しかし、本研究で用いた S P 質問は、新規高速船の端末交通手段を現在の交通手段から変更しない、という制約条件が課されてるため、個人パラメータに基づくセグメントは、その制約によって生じた S P データの選好の傾向に影響を受けている可能性があることが事例研究により明らかになり、S P・R P モデルの推定が行えなかった。個人パラメータによる方法は、S P という選好表示から個人の嗜好をダイレクトに求めるため、その有効性が高いことが事例研究からも確かめられたが、需要予測の観点から考えたとき、そのパラメータの将来分布の予測が困難で、個人属性などでは十分説明されにくいという報告 (森川, 白水 (1991)) もあるため、将来的なセグメントの帰属を予測することは難しいと考えられる。

潜在的態度変数に基づくセグメンテーションについては、S P モデルの結果から、個人の異質性を反映したセグメントを形成することができることを示した。また、S P・R P モデルに適用した結果から、各セグメント間の差異が、実際の行動結果を導入した分析にも有効であることを示し、潜在的態度変数によるアプローチの有効性を示すことができた。また、態度変数アプローチは線形構造方程式モデルを用いて態度変数を求めるため、その将来値の予測が可能であるという利点を有し、事例研究の結果からセグメントの有効性も確認された。しかし、潜在的態度変数によるアプローチで得られた、快適性重視セグメントについては、本研究では特にそのセグメントの特徴を表す快適性などの説明変数の導入ができなかった。今後、定性的な要因を重視するセグメントにおいても直接的に評価できるような変数を含めることによって、このアプローチの有効性を高めていく必要がある。また、統計的なクラスター分析で明確なセグメントを抽出できず、セグメントの意味付けが恣意的になってしまう可能性があることを示した。よって、今後、意味付を主眼にしたセグメンテーションと、統計的観点から行う 2 つのアプリオリセグメンテーションの特質を、マイクロシミュレーションなどを用いて政策分析を行い、内面的、外面的妥当性より有効性の確認が望まれる。



複数S P データとR P データおよび潜在的態度指標を用いて潜在セグメント分析を行った結果、統計的に有意なセグメントを抽出することができ、モデルの適合度も向上している。また、R P モードダミーの導入は、前に行った態度変数に基づいたアプリオリセグメントと比較してその有効性が薄れている。これは潜在セグメントモデルには系列相関項を導入したことによるものと考えられ、S P データの採取方法による選好の傾向は系列相関の考慮によって考慮できる可能性を示した。以上のことから、潜在セグメント分析が複数S P データの分析にも有効であることを示し、個人の嗜好の違いに基づいた潜在セグメント分析の適用の範囲を広げることができた。

以上本章では個人の異質性を考慮した非集計モデルシステムの構築をめざし、同じデータを用いて3種類の異なるアプローチを行った。その結果それぞれの利点・問題点が明らかになり、今後の研究発展の基礎を作ることに成功したと考える。

### 第3章 参考文献

大澤豊、片平秀貴、野本成明（1980）消費者研究における単調変換法を用いたコンジョイント測定法の応用に関する問題点、大阪大学経済学、Vol.30, pp.243-262.

小川孔輔（1981）「コンジョイント尺度」を与える最尤推定量について、経営志林、Vol.18, pp.37-52.

片平秀貴（1987-1）消費者研究とロジットモデル、鈴木雪夫・竹内啓編、社会科学の計量分析第7章、東京大学出版会

片平秀貴（1987-2）マーケティングサイエンス、東京大学出版会。

片平秀貴（1991）新しい消費者分析-LOGNAPの理論と応用、東京大学出版会、1991.

高田一尚、湯沢昭（1988）コンジョイント分析による個人行動モデルに関する研究、土木計画学研究・講演集、No.11, pp.707-714.

田中克明（1992）マーケットセグメンテーション、大澤豊ら編、マーケティングと消費者行動、第5章、有斐閣、pp.111-128.

西井和夫、北村隆一、近藤勝直、弦間俊彦（1995）観測されていない異質性を考慮した繰り返しデータに関するパラメータ推定法：Mass Point Model と Mixing Distribution Model、土木学会論文集、No.506/IV-26, pp.25-34.

藤原章正（1993）交通機関選択モデル構築における選好意識データの信頼性に関する研究、第4章、広島大学学位論文。

藤原章正、杉恵頼寧、張峻屹（1993）Mass Point 手法による交通機関選好モデルの消耗及び回答バイアスの修正、土木計画学研究・論文集、No.13, pp.587-594.

森川高行（1991）ステイティッド・プリファレンス、データの交通需要予測モデルへの適用に関する整理と展望、土木学会論文集、No.413/IV-12, pp.9-18.

森川高行、白水靖朗（1991）S P データを用いた交通需要予測のためのマーケットセグメンテーションに関する研究、土木計画学研究・講演集、No.14（1）、pp.589-596.

森川高行、Ben-Akiva M.（1992）R P データとS P データを同時に用いた非集計行動モデルの推定法、交通工学、Vol.27, NO4.

柳井晴夫、繁桝算夫、前川眞一、市川雅教（1990）因子分析、その理論と方法、朝倉書店、第5章。

湯沢昭、伊藤卓（1990）コンジョイント分析を利用した企業立地選好モデルの開発、土木計画学研究・講演集、No.13, pp.15-21.

Amemiya, T.（1985）Advanced Econometrics, Harvard University Press.

Andersen, E. B.（1982）Latent Structure Analysis: A Survey, Scandinavian Journal of Statistics, No.9, pp.1-12.

Ben-Akiva, M. and Morikawa, T.（1990）Estimation of Mode Switching Models from Revealed Preference and Stated Intentions, Transportation Research, Vol. 24 (A), pp.485-495.

Ben-Akiva, M., Morikawa, T. and Shiroishi, F.（1991）Analysis of the Reliability of Preference Ranking



- Data, *Journal of Business Research*, No.23, pp.253-268.
- Bergan, J. R. (1983) Latent-class Models in Educational Research, In E. W. Gordon (ed.), *Review of Research in Education 10*, American Education Research Association.
- Bhat, C. R. (1996) A Hazard-based Duration Model of Shopping Activity with Nonparametric baseline Specification and Nonparametric Control for Unobserved Heterogeneity, *Transportation Research*, Vol.30 (B), No.3, pp.189-207.
- Bolduc, D. and Ben-Akiva, M. (1990) A Multinomial Probit Formulation for Large Choice Sets, *International Conference on Travel Behavior*, vol2, pp243-258.
- Clogg, C. C. (1981) New Developments in Latent Structure Analysis, In D. J. Jackson and E. F. Borgatta (eds.), *Factor Analysis and Measurement in Sociological Research*, 8, pp.287-301.
- Currim, I. S. (1981) Using Segmentation Approaches for Better Prediction and Understanding from Consumer Mode Choice Models, *Journal of marketing Research*, XVIII, pp.301-309.
- Daganzo, C. (1979) *Multinomial Probit, The Theory and Application to Demand Forecasting*, Academic Press.
- Dillon, W. R. and Kumar, A. (1994) Latent Structure and Other Mixture Models, In R. P. Bagozzi (ed.), *Advanced Methods of Marketing Research*, Blackwell, pp.295-351.
- Fisher, G. W. and Nagin, D. (1981) Random versus Fixed Coefficient Quantal Choice Models, In Manski C. F. and McFadden D. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, pp.273-304.
- Formann, A. K. (1982) Linear Logistic Latent Class Analysis, *Biometrical Journal*, No.24, pp.171-190.
- Gopinath, D. A. and Ben-Akiva, M. (1996) Estimation of Randomly Distributed Value of Time, Prepared for 3rd Choice Symposium, New York.
- Gupta, S. (1991) Stochastic Models of Interpurchase Time with Time-Dependent Covariates, *Journal of Marketing Research*, 28, pp.1-15.
- Hensher, D. A. and Wrigley, N. (1986) Statistical Modeling of Discrete Choices in Discrete Time with Panel Data, *Behavioural Research for Transportation Policy*, VNU Science Press, pp.97-116.
- Kamakura, W. A. and Russel, G. J. (1989) A Probabilistic Choice Model for Market Segmentation and Elasticity Structure, *Journal of Marketing Research*, Vol. XXVI, pp.379-390.
- Kamakura, W. A. and Wedel, M. (1995) Life-Style Segmentation with Tailored Interviewing, *Journal of Marketing Research*, Vol. XXXII, pp.308-317.
- Kamakura, W. A., Kim, B. and Lee, J. (1996) Modeling Preference and Structural Heterogeneity in Consumer Choice, Prepared for 3rd Choice Symposium, New York.
- Kruskal, J. B. (1965) Analysis of Factorial Experiments by Estimating Monotone Transformation of the Data, *Journal of the Royal Statistical Society, Series B-27*, pp.251-263.
- Lam, S. H. and Mahmassani, H. S. (1990) Multinomial Probit Estimation: Computational Procedures and application *International Conference on Travel Behavior*, vol2, pp.228-242.
- Louviere, J. J. (1994) Conjoint Analysis, Bagozzi R. P. ed., 7th Chapter of *Advanced Methods of Marketing Research*, Blackwell, pp.223-259.
- Luce, D. (1959) *Individual Choice Behavior*, John Wiley and Sons.
- McFadden, D. (1989) A Method of Simulated Moments for Estimation of Multinomial Probit without Numerical Integration, *Econometrica*, 57.
- Morikawa, T., Ben-Akiva, M. and Yamada, K. (1992) Estimation of Mode Choice Models with Serially Correlated RP and SP Data, Presented Paper at 6th WCTR, Lyon.
- Morikawa, T. (1994) Correlating State Dependence and Serial Correlation in the RP/SP Combined Estimation Method, *Transportation*, 21, pp.153-165.
- Nicholas, T. L. (1993) *Random Coefficient Models*, Oxford Science Publications, Oxford.
- Ogawa, K. (1987) An Approach to Simultaneous Estimation and Segmentation in Conjoint Analysis, *Marketing Science*, Vol.6, No.1, pp.66-81.
- Srinivasan, V. and Shocker, A. D. (1973) Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgements, *Psychometrika*, Vol.38, No.4, pp.473-493/
- Swait, J. (1993) A Structural Equation Model of Latent Segmentation and Product Choice for Cross-Sectional Revealed Preference Choice Data, Presented Paper at the 1993 American Marketing Association Advanced Research Techniques Forum.
- Wedel, M., Kamakura, W. A., Desarbo, W. S. and Hofstede, F. T. (1995) Implications for Asymmetry, Nonproportionality, and Heterogeneity in Brand Switching from Price-wise Exponential Mixture Hazard Models, *Journal of Marketing Research*, Vol. XXXII, pp.457-462.
- Wishart, D. (1969) Mode Analysis: A Generalization of Nearest Neighbour Which Reduces Chaining Effects, In A. J. Cole, *Numerical taxonomy*, Academic Press, London, England, pp.282-319.

## 第4章 個人の異質性を考慮したパネル分析

### 4-1 本章の背景と目的

1章で述べたように、非集計行動モデルは様々な改良のための研究が行われ、多くの実務的成果を挙げている反面、当初期待されていたほどあらゆる局面に適用可能ではないことも明らかになりつつある。その原因として考えられていることの一つに、一断面のクロスセクションデータを用いることの問題が指摘されている。1988年のオックスフォード大学での国際会議（Oxford Conference on Travel and Transportation）では、行動のメカニズムに時間軸を考慮したダイナミック分析の必要性が主要な議題として取りあげられ、交通行動分析のみならず交通流配分理論などへの時間軸の導入が、主要な研究課題として浮かび上がってきた。そこで、交通行動分析の分野で近年その適用例が増えているのが、同一の個人や世帯などに対し、時点を変えて繰り返し観測したデータ（パネルデータ）の利用である。一時点の横断的なデータがクロスセクショナルデータと呼ばれるのに対し、パネルデータは、「母集団から抽出されたサンプルの全個人（パネルと呼ぶ）に、時点を変えて2回以上、面接調査して得られる情報」（Markas（1979））と定義される。

パネルデータによる交通行動のモデリングの利点は、同一個人に複数時点で尋ねたデータを用いることより、時間的に定常な非観測要因のコントロールが容易であることがあげられる。その結果、クロスセクショナルデータより小さいサンプルサイズを用いて同程度のモデル精度を確保できることと、交通行動の変化が明確に把握できる点にある（Kitamura（1990））。また、パネルデータを用いた特徴として、環境の変化に対する行動の変動を直接的に計測できるため、個人の選択行動の異質性の計測に適しているという特徴もある。これは、パネルデータが一人の個人に対して複数断面で調査を行うため、一人あたりの情報量が通常の一断面のクロスセクショナルデータと比較して多く、クロスセクショナルデータが交通条件の個人間の差で行動を説明するのにに対し、パネルデータは個人や交通条件の変化によって行動を分析することを可能にするため、各断面間で交通条件が変化した場合などは、個人から得られる情報量は格段に増えて行く。この結果、交通条件の変化に対して行動の変化が明確に測定でき、結果として、同一の状況にある個人が同一の状況変化に対して異なる反応を示した場合には、個人の異質性を明確に計測できると考えられる。たとえば、同一ODの経路選択を同一個人について継続的に調査した場合を考える。所要時間だけが変化したときの個人の経路選択の変化を数多く観測することで、個人の所要時間

に対する効用パラメータを計測することができる。そこで、同じ現象を多くのサンプルについて観測した場合に、所要時間時間に対する個人ごとの効用パラメータの違いが観測できることになる。パネルデータは、このような嗜好に関する個人の特徴、つまり効用パラメータの違いの計測が容易なデータであるため、パネルデータを3章で行った個人の異質性を考慮した交通行動分析に用いることは、非常に有効であると言える。

特に甲府パネルデータ（西井ら 1992）は、休日の買物場所選択を取り扱ったデータであるため、選択の自由度が高く、個人の嗜好の違いが顕著に現れてくると考えられる。このような非日常交通の行動分析は、嗜好の異質性を考慮することが非常に有効であると考えられる。そこで、本章でパネルデータを用いた、個人の異質性を考慮した非日常交通の行動モデル分析の事例研究を行う。具体的には、3章で展開した個人の異質性を考慮する手法をパネルデータにも適用し、異質性をより明らかにした動的な選択モデルの開発を試みるものである。本章での事例研究に用いる甲府パネルデータは、山梨県の甲府市を中心とする都市圏で行われている、休日の買物行動に関する調査である。休日の買物行動は、通勤・通学交通や平日の買物行動と比較すると、時間的制約や空間的制約等の制約が緩く、目的地選択や手段選択の選択肢集合が与えにくいこと、及びこれまでは交通計画上の重要性も低かったことから分析対象としてあまり取り上げられてこなかった。しかし、ライフスタイルの変化に伴い、休日交通が道路渋滞対策上重要となってきたうえ、制約がほとんど存在しないということは、逆に言えば個人の純粋な効用最大化行動が行われていると考えられ、パラメータに異質性が現れやすく、個人の異質性分析に適したデータであると言える。

また、大規模店舗法の改正などの規制緩和の流れに沿った形で、地方中核都市部郊外や地方中核都市の市街地周辺部に、自家用車での来訪を前提とした、買物、食事、娯楽といった諸活動サービスを提供するSC（ショッピング・コンプレックス）の立地が増加している。このような商業施設の大規模化・郊外化の流れの中で、今後休日の買物・娯楽交通はその交通計画上の重要性をますます高めていくものと思われる。このような時代背景の中で、個人の嗜好の違いを考慮した休日の買物行動をモデル化することは、どのような個人がどのような判断基準によって休日の買物目的地選択を行っているのかを示すことを可能にする。これは、休日の交通計画だけでなく、商業施設計画や、先に述べた商業施設の構造変化の流れの中で、地方都市で特に課題となっている、市街地中心部の商店街の相対的地盤沈下に対する振興政策などにも有益な情報を与え

ることができ、地方都市の都市計画の上でも非常に有益であると思われる。

## 4-2 パネルデータの特徴とパネルデータを用いた研究のレビュー

### 4-2-1 パネルデータの特徴

パネルデータは、繰り返し調査により得られるデータであり、その多断面性による情報量の多さからくる利点と、繰り返し調査を行うことによる欠点の両方を持ち合わせている。パネルデータの利点を示す例として以下のようなものがある（Hsiao（1986））。

「ある断面調査で、既婚女性の50%が就労していることが判明した場合、以下の2つの解釈も可能である。ひとつは、既婚女性は人生の半分就労し残り半分は就労しないというもので、もう一つの解釈は50%の既婚女性は常に働き、残り50%の既婚女性は常に働かないというものである。」

この例からわかるように、一断面のクロスセクショナルデータはこのような極端な2つの解釈が可能である。これらは明らかに間違った解釈であるが、クロスセクショナルデータによる集計からは、これを否定する根拠を明確に示すことは難しい。これら2つのモデルを、真のモデルと識別するためには、時間軸に沿った連続的な調査が必要である。この例でもわかるように、クロスセクショナルデータよりもパネルデータが優れている点は、時間的な変化を含むことによって、個人の差異によって行動の違いを説明するのではなく、個人の変化によって行動を分析することを可能にする点である。以下に交通行動の分析にパネルデータを用いた場合の特徴を簡単にまとめる（Kitamura（1990）、Hsiao（1986））。

- 1) 交通システムの変化により影響を受ける個人を含んだサンプルに、そのシステムの変化前後で繰り返し調査することで、個人の行動の変化を直接把握できる。
- 2) 行動の変化とその要因の変化を直接観察できるので、原因と結果の関係を導きやすい。
- 3) 同一人物を対象に調査を行うので、多時点クロスセクショナルデータと比較して、サンプルエラーを減らすことができ、多時点クロスセクショナルデータよりも小さなサンプルサイズで、同程度の精度のモデルが得られる。
- 4) データの収集点が増えるため、データの自由度が上がり、結果としてモデルの推定値の有効



性を高める。

- 5) 交通行動の各時点での状況を見ることができ、個人の履歴や経験のデータを得ることができ、行動の時間的依存性や状態依存性を明らかにできる。
- 6) 同一個人から多量の情報を得るため、個人の嗜好の違いに関する情報が得られ易い。
- 7) パネルサンプルが転居、死亡、あるいは調査拒否によってパネル調査から脱落していき、サンプル数が調査回数を増やすほど減少していく。
- 8) 繰り返し調査を行うことによって、データの精度が低下するパネル疲労の問題。
- 9) 複数回の回答行動が、被験者の思考回路に影響を与えるパネルコンディショニングの問題。
- 10) パネル調査を同一サンプルに対して行う場合、各サンプルの属性の変化によって、サンプルが母集団を代表しなくなるという問題。

ここで、パネルデータが原因と結果の関係を明確にできるという利点の、具体的な例を考えてみる。例えば、大都市圏域で交通行動を調査したところ、ある特定の地域の自家用車分担率が、他の地域に比べて有意に高かったとする。この意味はどのように解釈できるであろうか。これも先の既婚女性の就労調査の例のように、2通りの解釈が考えられる。それはある地域に住むと自家用車の利用率が高まるという解釈であり、もう一つは、自家用車の利用率が高い人がその地域に住んでいる、というものである。クロスセクショナルデータでは、このような原因と結果の関係を分析的に識別することはできないが、パネルデータを用いると、その地域に転出入した人を調査することで、原因と結果の関係を明確にすることができる。

#### 4-2-2 パネルデータを用いた既存研究のレビュー

個人の異質性に関する研究のレビューは3章で行ったため、本章ではパネルデータを用いた研究のレビューを行う。パネルデータの特質については4-2-1で述べたが、モデル論的観点からのパネルデータを用いたこれまでの研究は、先に挙げたような欠点を修正する指向のものが比較的多いことが言える。特に5章で取り扱う消耗バイアスに関する研究は、各国で数多くの研究者が様々な手法を用いて行われている。消耗バイアスの問題についての研究については5章で述べることとし、ここでは消耗バイアスに関する研究を除いた、パネルデータを用いた研究につ

いて簡単にレビューを行う。

Amemiya (1985) はパネルデータの各時点を独立として扱うとき問題となることは、個人の非観測異質性と真の状態依存であると指摘している。つまり、パネルデータを用いた場合には、一般に仮定される

$$P(y_{it} = 1) = F(x_{it}'\beta) \text{ for } i = 1, 2, \dots, n \quad (4.1)$$

$y_{it}$  : 個人  $i$  の  $t$  時点での選択結果

$x_{it}$  : 個人  $i$  の  $t$  時点での説明変数ベクトル

のような選択確率は不十分であり、これは任意の分布関数  $F$  に対して固定した  $t$  時点での周辺確率しか示せず、全時点での同時選択確率を示すものではないということである。たとえば、これの時点ごとの確率の積をとって全時点での同時選択確率とすることは、前の時点とは独立に各時点での確率が定義されていることとなり、4-1で述べた客観的事実に反する。Heckman and Willis (1977) は個人の非観測異質性を考慮するため (4.1) 式を次のように変形している。

$$P(y_{it} = 1 | u_i) = F(x_{it}'\beta + u_i) \text{ for } i = 1, 2, \dots, n \quad t = 1, 2, \dots, T \quad (4.2)$$

その上で被説明変数  $y_{it}$  を、非観測異質性項  $u_i$  が与えられたという条件付きで時間的に独立であると仮定している。Daganzo and Sheffi (1982) はパネルデータを、個人の離散的な選択を分析の単位とする非集計離散型選択モデルに適用する際には以下のような2つの問題を考慮しなければならないと述べている。一つは、ある時点の行動がその一時点以上前の行動結果に影響される「状態依存性 (State Dependence)」で、もう一点は同一個人の多時点の誤差項が互いに相関を持つ「系列相関 (Serial Correlation)」である。Heckman (1981) や Hensher and Wrigley (1986) は「状態依存性」を「真の状態依存 (True State Dependence)」, 系列相関を「疑似状態依存 (Spurious State Dependence)」であると定義し、パネル分析などの個人多時点データ分析の課題は、この2つの"State Dependence"を分離することであることを指摘している。ここに前者の例として、前時点での行動結果のみが異なり、他の点については全く同一の二人の個人について、次の時点での行動結果が異なるケースが挙げられる。この場合、ある時点における個人の行動を規定する要因の中に、それよりも前の時点の行動結果が明らかに含まれることを示し、それを真の"State Dependence"であるとしたのである。これに対し、疑似の"State Dependence"つまり系列相関は、両時点のモデルに共通して省かれた説明変数 (これを Omitted Variables と呼ぶ) や、モデルで表せ



ていない個人の嗜好の異質性などの影響によって、時点間の行動結果が相関を持っているように観測されることを指す。先の Amemiya (1985) が指摘した非観測異質性とここでの系列相関は、いずれも同様の原因によるもので、ほぼ同じ意味である。このように状態依存性があるデータや、誤差項間に相関があるデータを、独立と仮定して推定されたモデルは、そのパラメータの推定値にバイアスが生じる (佐和 (1970))。このような関係は、3章で用いた RP データと SP データの間にも見られる特徴である。このことは、消耗バイアスを持ったパネルデータにも、選択行動とパネル残留に同じ非観測要因が影響を与えている場合、各誤差項には同じ要因が影響するため、誤差項が独立で同一の分布をするというロジットモデルの前提に無理が生じてしまい、消耗バイアスと系列相関は同種の問題に帰着できる可能性がある。このような時点間の誤差相関を考慮して、モデル内に時点間の影響を取り込んだ動的行動モデルは、これまで様々に検討されており、Hensher and Wrigley (1986)、兵藤 (1988)、山田 (1991) の分類によると以下の5系統が存在する。

#### 1) マルコフ連鎖モデル (Amemiya (1985))

状態依存性を明示的にとり扱うモデルの代表例として、古典的なマルコフ連鎖モデル (Markovian Model) を挙げる。マルコフ連鎖モデルの仮定は「時点  $t$  の行動結果は時点  $t-n$  までの行動結果のみに依存する」というものである。このモデルは非常に単純明快なモデルであるが、ある時点の行動がすべて過去の行動結果により規定され、現時点での代替案の属性を全く説明変数に持たない点が、需要予測を行う実務上、非現実的であると考えられ、過去の行動履歴と現時点の属性を説明変数に持つ准マルコフモデルも提案されている。このモデルは現在の行動を規定する効用関数の説明変数に、現時点の属性と数時点前までの行動結果を含み、前述のマルコフ連鎖モデルの欠点を解消した形式を取っている。しかしながら、このようなモデルは説明変数が多くなるため、どの時点までの行動を説明変数として用いるかを理論的に明確に規定する必要があるが、これを決定する根拠に乏しく、一時点前のみを説明変数に用いる場合も、連鎖モデルであるため、最初の時点での選択は外生的に与えざるを得ない。似た考え方に基づくモデルとして、現在や過去の状態の存続期間を説明変数とする生存時間型 (Duration 型) のモデルがある。

#### 2) 過去の行動履歴を効用差で表現するモデル

1) で示した手法は、ダミー変数として過去の行動履歴を取り込んだモデルであった。次に

示す手法はそれらの一般形である。Daganzo and Sheffi (1982) はプロビット・タイプのランダム係数モデル (Random Coefficient Model) に、兵藤 (1988) は多項選択モデルの適用が可能なロジット・モデルにその適用例を示した。このモデルは一時点前の選択結果を説明変数とする准マルコフモデルの説明変数を、離散的なものから選択確率に応じて連続変数に置き換えたもので、 $t-1$  時点における選択肢  $i$  と  $j$  の効用関数の確定項の差を  $t$  時点における選択肢  $i$  の効用関数に組み込んでいる。

#### 3) 誤差項の共分散を考慮したモデル

前出の Daganzo and Sheffi (1982) は、系列相関をランダム係数モデル (Random Coefficient Model) の2項選択プロビットモデル (Binary Probit Model)、および3項選択プロビットモデル (Trinomial Probit Model) に適用した。これは、選択肢間、および時点間の誤差項に共分散を持たせることで、一般的なプロビットモデルの枠組みの中で扱うことができる。

#### 4) 誤差項に明示的に系列相関を取り込むモデル

Pas and Koppelman (1987) は旅行パターンの分析に、効用関数の誤差項をいくつかの部分に加法的に分解した、操作性に富む選択モデルを適用した。これは効用関数の誤差項を、時点を問わずに個人に共通な誤差、その時点においてすべての個人に共通の誤差と、真にランダムなホワイトノイズの三項に分離したものである。Kitamura and Bunch (1990) はこれと同様なモデルを自家用車保有行動分析に用いている。

#### 5) 事前行動結果によるセグメント

ある時点の行動結果を、その前の時点の選択結果の条件付き確率として定義するもので、兵藤 (1988) が提案している。具体的には全時点での選択結果により、セグメントした上での選択確率になる。現時点の選択は、前時点の選択に影響を受けるという仮定がされているが、系列相関や状態依存を構造的に扱っているわけではなく、異時点間の選択依存構造があらかじめ明確にされた場合、もしくは明確にする必要がない場合には有効であると考えられる。

#### 6) 個人内の説明変数の時点間の変化を用いるモデル

効用関数の説明変数を、個人内で時点間に変動のあったサンプルだけを用いるモデルで、説明変数に変化があったサンプルの効用は、2時点間の効用の差になっている。そのため、モード

ごとの定数項は消滅し、性別はもちろん年齢などの個人に固有の属性も、全員の変動が一緒なので説明変数に用いることはできない。この定式化により、時間で安定している個人の非観測異質性の影響は排除できる。ただし、変化がなかったサンプルは推定からはずさねばならず、その改善策として前の時点の選択を非変更者の効用として用いるモデルが提案されているが、その場合は先に述べた定常的な非観測異質性の排除ができなくなる。

Bradley (1992) は、ここに示したモデルを、同じデータを用いてそれぞれ推定し、シミュレーションを行って、パネルデータの特徴に応じたモデリング手法の提案を行っている。

このような基礎的なパネルデータの分析が続いた後、1992年にカリフォルニアで第1回の交通計画に関するパネル分析の会議が行われ、そこで最新の研究成果が発表された。その会議での研究発表は主に調査設計に関するもの、消耗分析に関するもの、調査とその分析の報告、活動の持続期間の分析であった。そこで発表された Hensher and Raimond (1992) のまとめた欧米での交通関連のパネル調査は20にもおよぶ。このように欧米ではパネル研究初期の3大プロジェクト (Sydney Automobile Panel, Cardiff Consumer Panel, Dutch National Mobility Panel) の後を受けて、個々の活動レベルの特徴把握に関する小規模のパネル調査も数多く行われているが、国内では、実務レベルでパネル調査が行われたことはほとんどなく、大学などの研究レベルでの試行にとどまっている。これは内田 (1993) の指摘によると、パーソントリップ調査などの大規模断面調査と比較して、パネル調査の特徴を生かす分析手法の確立が行われていないことと、交通行動モデルに対する交通計画側の需要がパネルデータを必要とするまでに至っていないことから、多くの費用を必要とするパネル調査が行政の側に避けられてきたためである。

パネルデータを用いて個人の異質性を考慮した研究として、Hensher and Wringley (1986) は、非観測異質性や状態依存について詳しく述べ、Cardiff Consumer Panel を用いて Mixing Distribution Model を推定し、非観測異質性を取り込んだ買物行動のパネル分析を行っている。同様に Dunn and Wringley (1985) も Cardiff Consumer Panel データを用いて Mixing Distribution Model を推定している。Hensher et. al. (1992) は自家用車の使用に関して、パネルデータの Left-censoring (Initial Condition) の異質性を考慮したモデルを構築している。ここでの Left-censoring の問題とは、行動の状態依存性を仮定した場合、一般に調査を開始する以前の状態についての情報は手に入れることが極めて困難であり、パネルデータを用いたとしても WAVE 1以降についてのみ状態依存を分析することが可能なだけであり、それ以前の状態による行動の依存性は分析できないという

問題を指す (Hensher and Wringley (1986))。Kitamura and Bunch (1989) ではパネルデータを用いて、状態依存性と非観測異質性の影響を考慮した自家用車の保有行動を分析している。Meurs (1993) は Dutch Mobility Panel を用いて、自家用車保有とその使用を、誤差項の相関を考慮した Mass Point Model を用いて分析している。Kim and Mannering (1992) は、活動継続時間を生存時間モデルを用いて分析し、その分布形をガンマ分布に従う個人の異質性項によって変動するワイブル分布と仮定し、Puget Sound Transportation Panel (Murakami and Ulberg (1992)) に適用している。

#### 4-3 パネルデータを用いた潜在セグメント分析

##### 4-3-1 モデルのフレームワーク

本章では、パネルデータを用いた個人の異質性を考慮した選択モデルの推定を目的としている。3章で指摘したとおり、収入、車の保有といった社会経済属性を効用関数に導入することで定数項を変動させる方法や、社会経済属性に基づいたアプリアリセグメントなどの手法は、同一セグメント内で効用パラメータが均質であるという仮定や、規定されたセグメントと効用パラメータの意味付けの根拠が不明確であるなどの問題があり、個人の嗜好を直接的に示す変数によるセグメンテーションや、セグメント内での効用パラメータの同一性の仮定を緩和したモデルの開発が必要である。それらの課題に対し、3章で個人パラメータや態度変数によるアプリアリセグメンテーション手法の提案や、態度変数を用いた潜在セグメントモデルの適用などを行い、その有効性を確認した。本章ではそれを受けて4-2で嗜好の異質性の検出に有効であると指摘したパネルデータを用いた目的地選択分析を行う。特に、3章でSP・RPデータへの適用でその有効性が確認された潜在セグメントモデルを主要な分析手法として採用し、パネルデータへの適用の有効性を検証する。パネルデータの動的特性については、4-2-2で示した誤差項に明示的に系列相関を考慮したモデルを構築する。3章では、潜在的態度変数を用いて潜在セグメントへの帰属度を表したが、ここでは、客観的な属性を潜在セグメントの帰属度に用いることとし、アプリアリセグメントとの比較を通じてその有効性を確認する。ここで、3章で示したように潜在セグメント分析手法の簡単な定式化を再び示す。

個人 $n$ をセグメント $s$ 、 $s=1,2,...,S$ に帰属させたときの効用関数を

$$U_{nls}^* = \beta_s X_{ni} + v_{nls} \quad (4.3)$$

とおく。ただし、

$U_{nls}^*$ ：セグメント  $s$  に帰属した場合の個人  $n$  の選択肢  $i$  に対する効用

$\beta_s$ ：セグメント  $s$  の効用パラメータベクトル

$X_{ni}$ ：個人  $n$  の属性と選択肢  $i$  の属性のベクトル

$v_{nls}$ ：セグメント  $s$  に属する個人  $n$  の選択肢  $i$  に対する効用の誤差項

このモデルは、各潜在セグメントごとに異なるパラメータを与えることによって、母集団内の効用パラメータの異質性を表している。ここで  $v_{nls}$  が独立で同一のガンベル分布に従うと仮定すると、セグメント  $s$  が与えられた条件付きの選択確率は

$$P_{ni|s} = \frac{e^{\beta_s X_{ni}}}{\sum_{j \in C_n} e^{\beta_s X_{nj}}} \quad (4.4)$$

と表される。

ただし、 $C_n$  は個人  $n$  の選択肢集合

母集団中の潜在セグメントの数が既知であるとするとき、各個人が潜在的なセグメントに属する確率を決定する変数として  $Y^*$  を仮定する。 $Y^*$  は各個人の態度指標などにより推定された潜在的態度変数と各個人の属性の線形和で与えられる。

$$Y_{ns}^* = \Gamma_G G_{ns}^* + \Gamma_z z_{ns} + \delta_{ns} \quad (4.5)$$

$Y_{ns}^*$ ：個人  $n$  のセグメント  $s$  に対する帰属度関数

$G_{ns}^*$ ：個人  $n$  の潜在セグメント  $s$  の帰属度に影響を与える態度変数

$z_n$ ：個人  $n$  の潜在セグメント  $s$  の帰属度に影響を与える社会経済属性ベクトル

$\Gamma_G, \Gamma_z$ ：未知パラメータベクトル

$\delta_{ns}$ ：個人  $n$  のセグメント  $s$  に対する帰属度関数の誤差項

このとき Luce の選択公理 (Luce (1959)) に従い、セグメント間の帰属比率が2つのセグメント間の帰属度の大小関係によって決定されると仮定する。 $\delta_{ns}$  が独立で同一のガンベル分布に従うとすると、各セグメントの帰属確率は次式で表されるような、ロジット型の確率になる。

$$W_{ns} = \frac{e^{\Gamma_G G_{ns}^* + \Gamma_z z_n}}{\sum_{k=1}^S e^{\Gamma_G G_{nk}^* + \Gamma_z z_n}} \quad (4.6)$$

意思決定者  $n$  が潜在セグメント  $s$  に帰属し、代替案  $i$  を選択する同時確率は、局所独立性の仮定により以下のような形で表すことが可能になる。

$$P_{nis} = P_{ni|s} W_{ns} \quad (4.7)$$

この条件付き確率を  $s$  に対して和を取ることにより、意思決定者  $n$  が代替案  $i$  を選択する周辺確率が得られる。

$$P_{ni} = \sum_{s=1}^S P_{ni|s} W_{ns} \quad (4.8)$$

ただし、 $S$  は潜在セグメントの総数

#### 4-3-2 用いるデータの概要

本研究では、先に述べたように、山梨県甲府市都市圏で行われた郊外に立地したショッピングコンプレックスへの来訪者を対象とした、甲府パネル調査を事例として分析を行う。この調査は、山梨大学の西井研究室が中心となって行っているもので、現在までに WAVE 7 (パネル調査では、調査時点のことを WAVE と呼ぶ) が整備された段階まで進んでいる (西井ら (1996))。甲府パネル調査は、S C の立地が市民の消費生活や交通行動に与える影響を分析するために開始された。調査は 1989 年秋に甲府市の郊外部に立地している「イーストモール S C バリオ」への来訪者を対象に、1 回目のアンケート調査がなされて以来、年 1 度の間隔で、休日における市民の買物・交通行動調査を行っている。具体的な内容は、調査日の一日の行動をダイアリー形式で尋ねたもの、回答者のライフサイクルステージや諸属性、休日の主な買物場所や、調査対象者の主な買物場所と思われる S C、甲府市中心街、最寄りのスーパーの 3 選択肢について、交通利便性指標と買物利便性指標の満足度、買物場所選択の理由がとられている。このデータの中から、本研究では事例研究で、以下にあげる 4 種類のデータを用いている。

1) 最近 1・2 カ月の休日に最もよく利用した買物場所 (選択結果)



- 2) 1) で回答した買物場所を利用した理由を、交通利便性と買物利便性に関する 10 項目の中から複数個選んだもの。(態度指標)
- 3) SC, 中心街, 最寄りスーパーそれぞれに対する交通利便性, および買物利便性を総合的に評価した 0 ~ 10 の 11 段階評点。(満足度評価)
- 4) 1) で回答した場所で, 最近行った買物の総購入金額などの買物行動に関する属性や, 世帯の社会経済属性。(客観的属性)

調査方法は, 甲府市郊外に立地している SC への来訪者に対して, 手渡しでアンケート票の配布を行い, 郵送で回収している。以降 1 回目と同様に SC におけるアンケート票の配布と並行して, 前回の調査まで回答のあった被験者に対し, その年に配布したものと同一アンケート票を郵送して回答を得ている。また, パネルサンプル以外の調査票配布は休日に行われている。これは, 休日は SC への来訪者が最も多く, SC の交通へ与える影響が大きいと考えられることと, 当日の移動を含めた行動全体を, アクティビティダイアリー形式で尋ねるため, 行動の自由度が大きく, 被験者の環境の変化が活動に与える影響が明確に現れやすい休日を対象としている。

#### 4-3-3 客観的属性を用いた潜在セグメント分析

ここで, (4.5) 式の潜在セグメントへの帰属度変数を客観的属性のみで規定するモデルを構築し, 同様の変数を用いたアプリアリセグメント手法と比較検討を行う。4-3-2 で述べた甲府パネル調査の WAVE 3 と WAVE 4 のデータを使って, 買物目的地選択を非集計モデルを用いて定式化した。選択肢は先に挙げた SC, 中心部商店街, 最寄りスーパーの 3 項とし, 効用関数は, 以下に示す選択肢別の変数と SC 定数項の線形和とした。

- 1) 自宅から各買物目的地への距離
- 2) 各買物目的地に対する回答者の交通利便性満足度 (11 段階評価)
- 3) 各買物目的地に対する回答者の買物利便性満足度 (11 段階評価)

これらのデータが総てそろっている WAVE 3, WAVE 4 の両方ともに参加しているパネルサンプルを用いた。

WAVE 3 の各選択肢の効用関数は以下の (4.9) 式から (4.11) 式のようになり, WAVE 4 についても同様に定義されている。

$$U_{n,SC}^{w3} = \beta Z_{n,SC}^{w3} + \epsilon_{n,SC}^{w3} \tag{4.9}$$

$$U_{n,C}^{w3} = \beta Z_{n,C}^{w3} + \epsilon_{n,C}^{w3} \tag{4.10}$$

$$U_{n,SP}^{w3} = \beta Z_{n,SP}^{w3} + \epsilon_{n,SP}^{w3} \tag{4.11}$$

ただし,

$U_{n,SC}^{w3}$ ,  $U_{n,C}^{w3}$ ,  $U_{n,SP}^{w3}$ : それぞれ個人  $n$  の SC, 甲府市中心街, 最寄りスーパーの効用

$Z_{n,SC}^{w3}$ ,  $Z_{n,C}^{w3}$ ,  $Z_{n,SP}^{w3}$ : 個人  $n$  の各選択肢に対する自宅からの距離, 交通利便性満足度, 買物利便性満足度のベクトル

$\beta$ : 未知パラメータベクトル

$\epsilon_{n,SC}^{w3}$ ,  $\epsilon_{n,C}^{w3}$ ,  $\epsilon_{n,SP}^{w3}$ : それぞれ個人  $n$  の各選択肢に対する効用の誤差項

この特定化に従って, 様々な個人属性に基づいてアプリアリセグメンテーションした買物目的地選択モデルの推定を試みた。その結果, 最も有効なセグメントを形成したのは, 当日の総購入金額によるものであった。その推定結果を表 4-1 に示す。当日の総購入金額を答えているパネルサンプルが少なかったため, 各 WAVE のサンプル数は 51 人, 総観測数にして 102 にとどまった。

| 表 4-1 購入金額によるアプリアリセグメントモデルの推定結果 |                  |                  |
|---------------------------------|------------------|------------------|
|                                 | 購入金額 7500 円以上※ 2 | 購入金額 7500 円未満※ 2 |
| SC 定数項 3 ※ 1                    | 1.12 (1.9)       | 0.0348 (0.1)     |
| SC 定数項 4 ※ 1                    | 0.633 (1.1)      | 0.0859 (0.2)     |
| 距離                              | -0.820 (-2.9)    | -0.561 (-3.1)    |
| 交通利便性                           | 0.139 (1.0)      | 0.309 (3.0)      |
| 買物利便性                           | 0.701 (2.8)      | 0.341 (2.2)      |
| 観測数                             | 42               | 60               |
| $\bar{p}^2$                     | 0.384            | 0.116            |
| $\bar{p}^2$                     | 0.185            |                  |

※ 1 表中の 3 および 4 は WAVE を表す ※ 2 ( ) 内は t 値



このモデルでは、当日の購入金額7500円を境として全サンプルを2つのグループに分け、それぞれセグメント内では同一のパラメータを持つと仮定したアプリオリセグメントである。当日の購入金額は内生的変数という考え方ができるが、それは購入金額が買物目的地と件での決定ということであり、この事例研究では、買物の金額（品目）があらかじめ条件として与えられた場合の目的地選択行動であるとの仮定をおき、買物金額を外生変数としてセグメント規定変数に採用した。このモデルの推定結果を見ると、購入金額の高い層は買物利便性を重視し、S Cに対して好意的な層であると考えられ、購入額の少ない層はS C定数項が有意でなく、買物利便性だけでなく交通利便性も重視している層であることがうかがえる。ただし、モデルのフィットは購入金額7500円以上のセグメントは良いが、7500円未満のセグメントではモデルのフィットがあまり良くない。これは購入金額が高い方のセグメントでは、定数項が有意であることによるものと考えらる。

続いて、同じ購入金額を用いて4-3-1で概説した潜在セグメントモデルを用いて買物目的地行動を分析する。(4.5)式で表されるセグメント帰属度を先のアプリオリセグメントと比較するため、以下のように特定化した。

$$Y_{ns}^* = \Gamma_{zs} pc_n + \delta_s \quad (4.12)$$

ただし

$pc_n$  : 個人  $n$  の総購入金額

(4.12)式で表されるセグメント帰属度と、(4.9)～(4.11)式で表された選択モデルを(4.8)式中に代入して得られる個人の選択確率を、個人の尤度として最尤推定法を用いて同時に推定した結果が表4-2である。表中のスケールパラメータとは、3章でS P・R Pモデルに適用したときと同様に、WAVE 3とWAVE 4の誤差項の分散の大きさの違いを示すパラメータであり、以下の関係が成立している。

$$\text{var}(\varepsilon^{\text{wave}3}) = \mu^2 \text{var}(\varepsilon^{\text{wave}4}) \quad (4.13)$$

表4-2の推定結果を見ると、潜在セグメントモデルは自由度修正済み尤度比が高くなり、的中率も向上している。帰属度モデルに用いた総購入金額は7500円からの偏差に変換しており、その推定値が正であることから、購入金額の多い人ほどセグメント1に帰属する確率が高くな

表4-2 総購入金額による潜在セグメントモデル

| 変数名         |           | セグメント1※2       | セグメント2※2      |
|-------------|-----------|----------------|---------------|
| 選           | S C定数項3※1 | 4.48 (1.1)     | -1.92 (-1.5)  |
| 択           | S C定数項4※1 | 2.32 (1.4)     | -1.44 (-1.4)  |
| モ           | 距離        | -1.52 (-1.4)   | -0.614 (-2.2) |
| デ           | 交通利便性     | 0.542 (1.5)    | 0.388 (1.5)   |
| ル           | 買物利便性     | -0.197 (-0.9)  | 1.42 (2.4)    |
| 帰属度モデル      |           | 総購入金額          |               |
|             |           | 0.0617 (1.3)   |               |
| スケールパラメータ   |           | 1.34 (0.43) ※3 |               |
| サンプル数       |           | 51             |               |
| $\bar{p}^2$ |           | 0.374          |               |

※1：表中の3および4はWAVEを表す ※2：( )内はt値 ※3：1を基準としたt値

り、購入金額が低い人ほどセグメント2に帰属する確率が高くなる。3章でも述べたように、潜在セグメントモデルは各個人がその帰属度に応じて、セグメント1と2のパラメータ値の内分点をとっていると解釈される。ここが、同一セグメント内での効用パラメータが均質と仮定をおくアプリオリセグメントと比較して、セグメント内では効用パラメータが同一という制約が緩和されて、個人パラメータアプローチに近い性質を持つこととなり、その優位性が現れる点になる。現に適合度は、アプリオリセグメントの全体の適合度と比較して著しく高くなった。これはアプリオリセグメントモデルは購入金額7500円未満のセグメントの適合度が低いため、モデル全体としての適合が悪くなったことを考慮しても、潜在セグメントモデルの有効性を示すことができた。推定値をそれぞれ見ていくと、セグメント1はS Cの定数項が大きいことと、交通利便性のパラメータが買物利便性のパラメータよりも相対的に大きく、交通利便性を重視していることがうかがえる。一方セグメント2は、S C定数項の値は負であって潜在的にS Cを好まず、どちらかという交通利便性よりも買物利便性を重視するという結果となっている。これを先のセグメント帰属モデルの解釈とあわせても、対照的で解釈が容易なセグメントが抽出されたと考えられる。

また比較のため、購入金額を効用関数の説明変数としたモデルを表4-3に示した。この結果を見ると、総てのパラメータは潜在セグメントモデルのセグメント1とセグメント2の中間の値を取っている。潜在セグメントモデルを用いた場合の各サンプルにとっての効用パラメータは、2つの潜在セグメントのパラメータを、各サンプルのセグメント帰属確率をウェイトとした加重平均になっていると考えることができるため、この推定結果は予想された通りである。また、潜在セグメントモデルを用いた場合と比較して、定数項を除いた各パラメータのt値は高い。今回の事例研究では用いたサンプルの数に対して、潜在セグメントモデルを用いた場合、推定したパラメータが多くなるため、潜在セグメントの各パラメータの有意性が低く、この結果によって効果を断定することは一定の危険を伴うが、効用関数に個人属性を導入するモデルや、アприオリセグメントを行ったモデルより、適合度、的中率ともに優れているため、潜在セグメントモデルの有効性を示すことができたと考えられる。

表4-3 購入金額を説明変数とするモデル

| 変数名            | 推定値※2          |
|----------------|----------------|
| SC定数項3※1       | 0.445 (1.0)    |
| SC定数項4※1       | 0.265 (0.5)    |
| 距離             | -0.759 (-3.4)  |
| 交通利便性          | 0.300 (2.4)    |
| 買物利便性          | 0.557 (2.8)    |
| 購入金額           | 0.106 (2.0)    |
| スケールパラメータ      | 0.835 (0.5) ※3 |
| サンプル数          | 51             |
| $\bar{\rho}^2$ | 0.292          |

※1：表中の3および4はWAVEを表す ※2：（）内はt値 ※3：1を基準としたt値

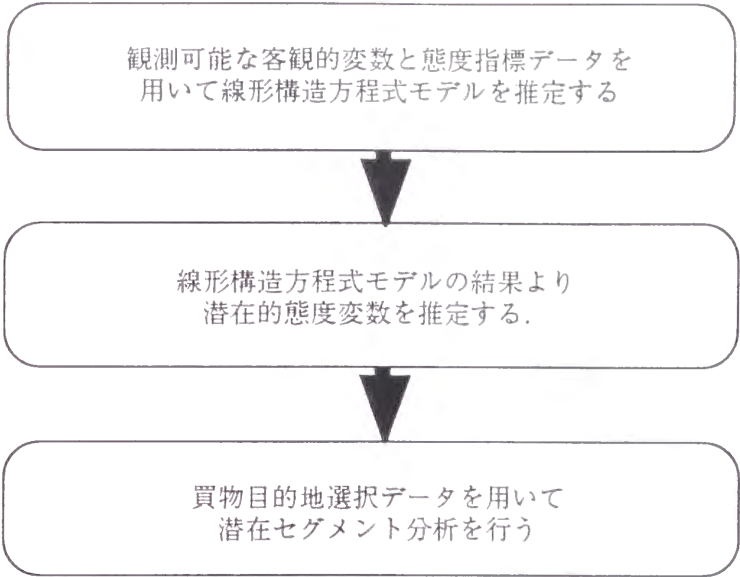


図4-1 潜在的態度を用いた潜在セグメント分析

4-3-4 潜在的態度を用いた潜在セグメント分析

甲府パネル調査では、4-3-1で述べたように買物目的地選択の理由を尋ねたデータが取られている。3章で展開した個人の異質性を考慮した選択モデル分析で行ったものと同様に、ここで得られた態度データを用いて（態度についての説明は3章を参照）、個人の潜在的な態度変数を算出し、潜在セグメントモデルの帰属度変数の説明変数として用いる。その方法論は3-4に示したものと同様でフローチャートで表すと図4-1のようになる。

ここでは、4-3-3で購入金額を $Y_{ns}^*$ と規定したのと同様に、(4.5)式で表されるセグメントへの帰属度 $Y_{ns}^*$ を、潜在的態度変数の線形和で表すものである。以下3-4と同様に、潜在的態度変数と観測可能な変数の関係を線形構造方程式モデルを用いて定める。

測定方程式

$$Y_n = \Lambda G_n^* + \varepsilon \tag{4.14}$$

構造方程式

$$G_n^* = Bw_n + \zeta \tag{4.15}$$

ただし、

$Y_n$  : 個人  $n$  の観測された態度指標ベクトル

$G_n^*$  : 個人  $n$  の潜在的な態度変数

$W_n$  : 態度変数に影響を与える個人  $n$  の社会経済属性ベクトル

$A, B$  : 未知パラメータ行列

$\epsilon, \varsigma$  : それぞれ多変量正規分布に従う誤差項

甲府パネルデータでは、態度指標として以下の 10 項目が得られている。

[ ( ) 内は後の定式化の時に用いる変数名 ]

- 1) 駐車場を探す時間や待ち時間の長さ (*time*)
- 2) 駐車場の停めやすさ (*ease*)
- 3) 駐車料金の安さ (*fare*)
- 4) 駐車場の位置のわかりやすさ (*plac*)
- 5) 駐車場までの経路のわかりやすさ (*rout*)
- 6) 商品の種類の豊富さ (*vari*)
- 7) 品質の良さ (*qual*)
- 8) 価格の安さ (*price*)
- 9) 広告などの情報量の多さ (*adve*)
- 10) 他の目的地への移動のし易さ (*loca*)

これらの態度指標は、「重視する」、「重視しない」のどちらかを選択した離散変数で得られ、これら 10 の態度指標に対する満足度がアンケート調査中で尋ねられている。式 (4.9) ~ (4.11) の効用関数の説明変数に用いた変数のうち、交通利便性満足度はこれら 10 個の態度指標のうち 1) ~ 5) の総合評価、買物利便性満足度は 6) ~ 10) の総合的評価という位置づけがある。このように観測された態度指標が交通利便性満足度に関連するものと、買物利便性満足度に関連するものになっているため、潜在的な態度要因として、買物利便性重視態度、交通利便性重視態度の二つを考え、(4.14) (4.15) 式で表される線形構造方程式モデルを以下のように特定化した。

構造方程式

$$G_l^* = \beta_{1l}age + \beta_{2l}fem + \beta_{3l}job + \beta_{4l}clis + \beta_{5l}com + \beta_{6l}pay + \beta_{7l}sbur + \varsigma_l \quad (4.16)$$

ただし、

$l$  : 1 ; 交通利便性重視変数, 2 ; 買物利便性重視変数

$age$  : 1 ; 40 歳以上, 0 ; その他

$fem$  : 1 ; 女性, 0 ; 男性

$job$  : 1 ; 就業者, 0 ; その他

$car$  : 家庭の自家用車保有台数

$clis$  : 免許保持者一人当たりの家庭の自家用車台数

$com$  : 1 ; 3 品目以上購入した, 0 ; その他

$pay$  : 総購入金額

$sbur$  : 甲府中心街から居住地までの距離

測定方程式

$$\begin{bmatrix} y_1(time) \\ y_1(ease) \\ y_1(fare) \\ y_1(plac) \\ y_1(rout) \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} G_1^* + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix} \quad (4.17)$$

$$\begin{bmatrix} y_1(vari) \\ y_1(qual) \\ y_1(pric) \\ y_1(adve) \\ y_1(loca) \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \\ \lambda_8 \end{bmatrix} G_2^* + \begin{bmatrix} \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \end{bmatrix} \quad (4.18)$$

(4.17) 式は交通利便性重視変数の測定方程式であり、(4.18) 式は買物利便性重視変数の測定方程式である。パラメータ同定のため、測定方程式のパラメータは各測定方程式ごとに一つづつ 1 に固定されている。この事例研究では、先に述べたようにセグメントを規定する潜在的態度変数を 2 つに設定したため、セグメント帰属変数の自由度は 2 となるので、潜在セグメントの数は最大で 2 となる。よって選択モデルの潜在セグメント数をあらかじめ 2 として、各セグメントへの帰属度を表す潜在変数を  $G_1^*$ ,  $G_2^*$  の態度変数で規定する。

$$Y_{ns}^* = \Gamma G_{ns}^* + \delta_s \quad s=1,2 \quad (4.19)$$

また、効用関数は (4.9) ~ (4.11) 式と同様に、自宅からの距離、交通利便性および買物利便性の満足度評価を用いて以下のように定式化した。



$$U_{ins}^* = \alpha_{1is} + \alpha_{2is} dist + \alpha_{3is} trav + \alpha_{4is} shav + v_{is} \quad s = 1, 2 \quad (4.20)$$

*dist* : 自宅から各目的地までの距離

shav: 買物利便性の満足度

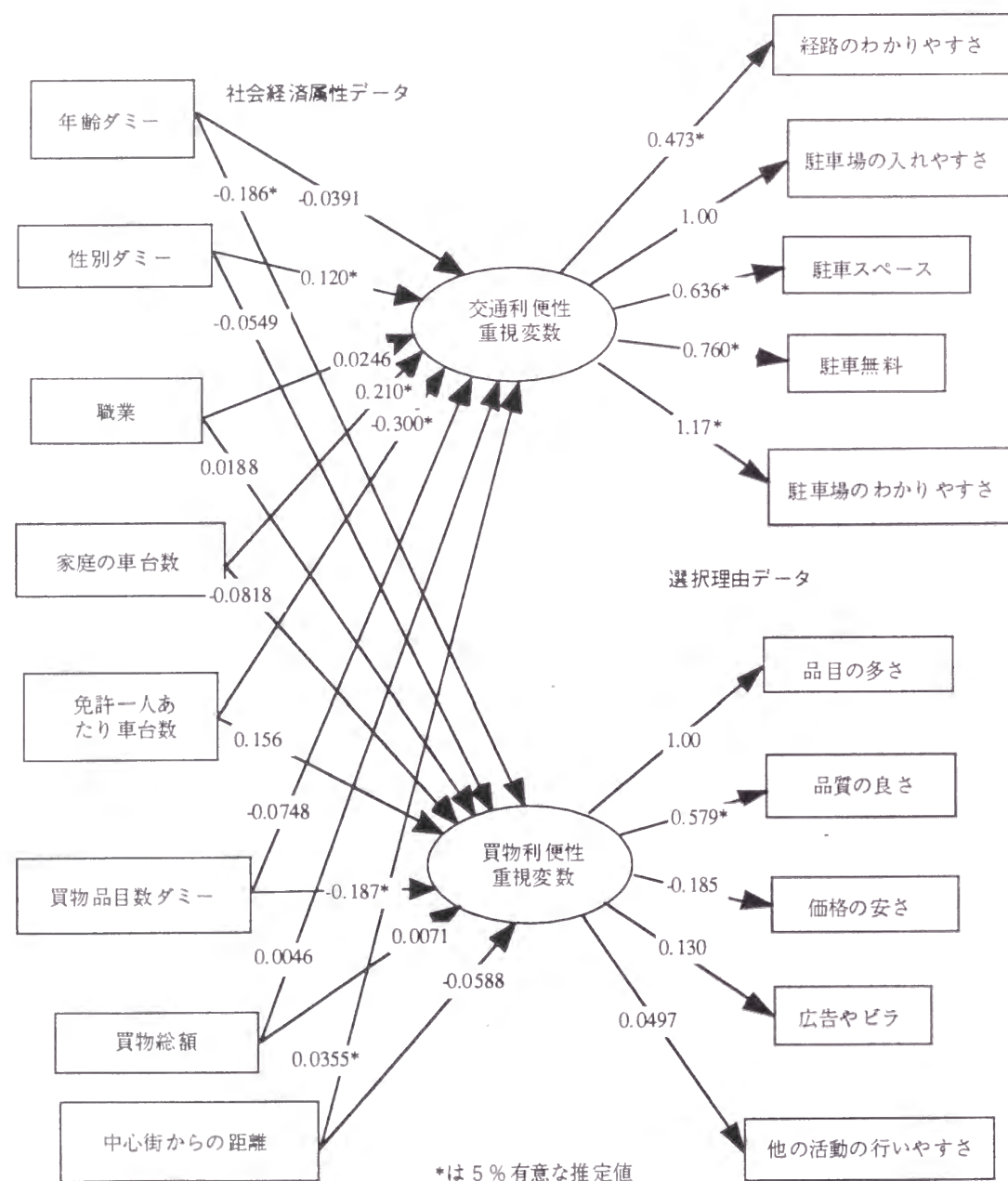


図 4-2 線形構造方程式モデルの推定結果

これらの定式化に基づいて、線形構造方程式モデルと潜在セグメントモデルを段階的に推定した。線形構造方程式モデルの推定結果を図4-2に示した。図中の各矢印に記されている数値は各パラメータの推定値である。この推定結果を見ると、測定方程式は買物利便性重視変数と価格の安さとの関係が負の符号となり、直感的な符号とは異なるが有意ではない。その他は正の符号を持ち、特に交通利便性重視変数では有意なものが多い。一方、構造方程式は中心街からの距離が増大すると交通利便性を重視するなど、妥当な符号で有意な変数が交通利便性重視変数には多い。買物利便性重視変数は、買物品目ダミーが負の符号で有意となるなど直感的基準とは異なるが、これは先に述べた測定方程式で、価格の安さが負の符号を持つことと関連し、品目の多さや品質の良さと価格の安さが同時に満たされることが少なく、トレードオフ関係を持っていることと、買物品目の多い人は価格を重視する傾向があると思われるため、このような推定結果が現れたと考えられる。以上のような考察より、これらはそれぞれ交通利便性重視変数と買物利便性重視変数をそれぞれ表していると考えることが可能であるとする。この推定結果をもとに、2章での定式化に基づいて(4.21)式によってそれぞれの態度変数の推計値(Fitted Value)を求め、(4.19)式の潜在態度に用いる。

$$\hat{G}_{\mu\nu}^* = \hat{B}\hat{w} + \hat{\Psi}\hat{\Lambda}'(\hat{\Lambda}\hat{\Psi}\hat{\Lambda}' + \hat{\Theta})(Y - \hat{\Lambda}\hat{B}\hat{w}) \quad (4.21)$$

セグメント 1 は相対的に交通利便性重視変数が帰属度を定めるセグメントであり、セグメント 2 は相対的に買物利便性重視変数が帰属度を定めるセグメントである。推定結果を見るとセグメント 1 では交通利便性評価の係数が正でかつ比較的に有意な値になっている。同様にセグメント 2 では買物利便性評価変数の係数が正の大きな値になっている。また定数項がセグメント 1 ではいずれも正の値になっているが、セグメント 2 ではいずれも負の値になっている。このことからセグメント 1 は潜在的に SC を好み交通利便性を重視するセグメントであり、セグメント 2 は潜在的に SC を好まず買物利便性を重視するセグメントであると考えられ、これらの組み合わせは直感的な予想に矛盾しない結果となっている。



表 4 - 4 態度変数による潜在セグメントモデルの推定結果

| 変数名            |               | セグメント 1 ※ 2    | セグメント 2 ※ 2   |
|----------------|---------------|----------------|---------------|
| 選              | S C 定数項 3 ※ 1 | 3.29 (1.8)     | -1.64 (-1.0)  |
| 択              | S C 定数項 4 ※ 1 | 1.90 (1.9)     | -1.25 (-1.0)  |
| モ              | 距離            | -1.29 (-1.8)   | -0.632 (-2.2) |
| デ              | 交通利便性         | 0.508 (1.7)    | 0.320 (1.4)   |
| ル              | 買物利便性         | -0.173 (-1.0)  | 1.46 (1.8)    |
| 帰属度モデル 態度変数    |               | 2.41 (1.3)     |               |
| スケールパラメータ      |               | 1.53 (0.6) ※ 3 |               |
| 観測数            |               | 1 0 2          |               |
| $\bar{\rho}^2$ |               | 0.261          |               |

※ 1 : 表中の 3 および 4 は WAVE を表す ※ 2 : ( ) 内は t 値 ※ 3 : 1 を基準とした t 値

系列的に安定しているものも多く、パネルデータなど同一個人から繰り返しデータを採取する場合には、確率項には意思決定者個人の系統的な誤差が各選択肢に含まれ、個人の各時点での確率項間に相関が生じることが多い。このような場合に、時点間で誤差項は独立であるという仮定に基づいて、同一個人からのデータも他人から得たデータと同じように扱ってモデルの推定を行うと、誤差項の系列相関の影響によって、推定したモデルのパラメータの分散値にはバイアスが生じてしまう。そこでこの相関によるバイアスを除去するために、3 章でも用いた Morikawa (1994) で提案されいてる、効用関数の誤差項を個人に時点を通じて共通の項  $\lambda$  と真にランダムなホワイトノイズを表す項  $v$  に分解し、時点間での誤差項の相関を解消する手法を用いる。以下、系列相関と潜在セグメントを考慮した選択モデルを定式化する。

セグメント  $s$  が与えられた条件のもとで  $t$  時点における効用関数

$$U_{nit|s} = V_{nit|s} + \varepsilon_{nit|s} \quad (4.22)$$

$t - 1$  時点における効用関数

$$U_{nit-1|s} = V_{nit-1|s} + \varepsilon_{nit-1|s} \quad (4.23)$$

$V_{nit|s}$ ,  $V_{nit-1|s}$  : 各時点での個人  $n$  の効用関数の確定項

$\varepsilon_{nit|s}$ ,  $\varepsilon_{nit-1|s}$  : 個人  $n$  のそれぞれの時点における代替案  $i$  についての誤差項

ここで、誤差項に明示的に系列相関を取り込むため、誤差項をそれぞれの代替案について各個人で共通な部分と、真にランダムな部分の二つに分離する。系列相関は選択肢固有変数の影響も含まれているため、各選択肢ごとに異なると考えられる。そこで、各選択肢ごとに系列相関項を導入する。

$$\varepsilon_{nit|s} = \lambda_{ni|s} + v_{nit|s} \quad (4.24)$$

$$\varepsilon_{nit-1|s} = \lambda_{ni|s} + v_{nit-1|s} \quad (4.25)$$

$\lambda_{ni|s}$  : セグメント  $s$  が与えられたもとでの個人  $n$  の代替案  $i$  に対する固有の誤差項

このように、時点間で個人ごとに共通の誤差項を導入することにより、二つの確率項を含む効用関数が定義され、真の誤差項である  $v$  の独立性が保たれる。この考えに基づいた時点  $t$  における効用関数は以下ようになる。

$$U_{nit|s} = V_{nit|s} + \lambda_{ni|s} + v_{nit|s} \quad (4.26)$$

さらに、4 - 3 - 3 と同様に、スケールパラメータ  $\mu$  を導入して効用関数のスケールを統一する。ここに、スケールパラメータ  $\mu$  は、

$$Var(v_{nit|s}) = \mu^2 \cdot Var(v_{nit-1|s}) \quad (4.27)$$

という関係を示し、効用関数の誤差項  $v$  は、時点に対し独立で同一な分散を持つ確率項と定義する。このとき  $t$  時点およびスケールパラメータを導入して再定義した  $t - 1$  時点の効用関数は以下ようになる。

$$U_{nit|s} = V_{nit|s} + \lambda_{ni|s} + v_{nit|s} \quad (4.28)$$

$$U_{nit-1|s} = \mu V_{nit-1|s} + \mu \lambda_{ni|s} + \mu v_{nit-1|s} \quad (4.29)$$

ここで注意しなければならないことは、系列相関項は各選択肢ごとに相対的な値を取るため、モデル推定のためには一つの選択肢に対する系列相関項をゼロに固定する必要がある。その上で、各選択肢に対する 2 つの誤差項のうち、 $v$  の分布形を定義することにより  $\lambda$  が既知であるときの

$t$  時点の条件付き選択確率  $P_{nt}(i|\lambda)$  が得られる。このとき、ある個人が  $t$  時点、 $t-1$  時点でそれぞれ  $i, j$  という代替案を選択する同時選択確率  $P_n$  は、 $\lambda$  の分布形を与えることにより次のように与えられる。

$$P_n = \int \int \cdots \int_{-\infty}^{+\infty} P_{nt}(i|\lambda) \cdot P_{nt-1}(j|\lambda) \cdot f(\lambda) d\lambda_{1n} d\lambda_{2n} \cdots d\lambda_{ln,n} \quad (4.30)$$

$f(\lambda)$  :  $\lambda$  の同時密度関数  
 $ln$  : 個人  $n$  の選択肢集合  $Cn$  に含まれる選択肢の数

系列相関を考慮したモデルのパラメータ推定は、(4.30) 式で与えられる個人の選択確率を個人の尤度として、最尤推定法により未知パラメータを推定する。尤度関数中の積分については、Gauss-Legendre の公式（渋谷（1965））に従って数値的に評価する。

このモデルに 4-3-1 で解説した潜在セグメントモデルを導入した場合、 $t$  時点において個人  $n$  がセグメント  $s$  に帰属する確率は、

$$W_{nst} = \frac{e^{(r_{Gi} G_{nst}^*)}}{\sum_{k=1}^S e^{(r_{Gk} G_{nst}^*)}} \quad (4.31)$$

潜在セグメントを考慮したときの、条件付き選択確率  $P_{nt}(i|\lambda)$  は (4.7) 式より、

$$P_{nt}(i|\lambda) = P_{nt|s}(i|\lambda) W_{nst} \quad (4.32)$$

となる。ある個人が  $t$  時点、 $t-1$  時点でそれぞれ  $i, j$  という代替案を選択する同時選択確率  $P_n$  は、 $\lambda$  の分布形を与えることにより次のように与えられる。

$$P_n = \int \int \cdots \int_{-\infty}^{+\infty} P_{nt}(i|\lambda) \cdot P_{nt-1}(j|\lambda) \cdot f(\lambda) d\lambda_{1n} d\lambda_{2n} \cdots d\lambda_{ln,n} \quad (4.33)$$

$f(\lambda)$  :  $\lambda$  の同時密度関数

本研究では、 $v$  に I.I.D. ガンベル分布、 $\lambda$  に独立な標準正規分布を仮定し、これを個人の尤度として、(4.30) 式と同様に最尤推定法を用いて未知パラメータを推定する。その推定結果を表 4-5 に示す。

表 4-5 系列相関を考慮した態度変数による潜在セグメントモデルの推定結果

| 変数名            |               | セグメント 1 ※ 2    | セグメント 2 ※ 2   |
|----------------|---------------|----------------|---------------|
| 選              | S C 定数項 3 ※ 1 | 4.40 (1.7)     | -1.99 (-1.3)  |
| 択              | S C 定数項 4 ※ 1 | 2.34 (1.8)     | -1.29 (-1.1)  |
| モ              | 距離            | -1.56 (-2.0)   | -0.732 (-2.8) |
| デ              | 交通利便性         | 0.536 (1.8)    | 0.403 (1.7)   |
| ル              | 買物利便性         | -0.177 (-1.0)  | 1.53 (2.2)    |
| 帰属度モデル         |               | 2.57 (1.6)     |               |
| スケールパラメータ      |               | 1.35 (0.8) ※ 3 |               |
| 観測数            |               | 1 0 2          |               |
| $\bar{\rho}^2$ |               | 0.255          |               |

※ 1 : 表中の 3 および 4 は WAVE を表す ※ 2 : ( ) 内は t 値 ※ 3 : 1 を基準とした t 値

系列相関項は、定数項が個人ごとに同一の分布をしていると解釈することも可能であるため、個人の嗜好の違いの要素を持っている。その影響により得られるセグメントの特性は異なるかと思われたが、表 4-4 に示した系列相関を考慮していないモデルと比較して推定値に大きな違いが見られず、ほぼ同様の解釈が可能なパラメータを持つセグメントが得られた。系列相関項を導入したことが、定数項の推定値の WAVE 間での差異にそれほど影響を与えていない事を勘案すると、潜在的な異質性は、WAVE 間での定数項の異質性と、2つの潜在的なクラスを仮定することで、十分であるのかも知れない。この2つのモデルの推定結果を詳しく比較すると、系列相関を考慮したモデルでは各パラメータの絶対値が系列相関を考慮しないモデルより大きくなっており、効用関数の真の誤差項の分散が小さくなっていることを示していると考えられる。また、適合度および的中率は系列相関を考慮したことにより低下した。これは系列相関が誤差項に存在する場合、誤差項全体の分散は過小推定されてしまうことがあるため、モデルの適合度が不当に高く評価されてしまうためと考えられる。

#### 4-4 本章のまとめ

以上事例研究で得られた結果をまとめると以下の4点になる。

1) パネルデータに客観的変数(購入金額)をセグメント規定変数とする潜在セグメントモデルを適用し、アプリアリセグメントや効用関数への変数導入手法と比較した結果、非常に高い適合度と、的中率を得ることができ、かつ抽出されたセグメントはセグメント帰属変数と関連を持ち、解釈が容易なものであった。

2) 比較的容易に得ることのできる態度指標と個人属性を用いて、線形構造方程式モデルを推定したところ、モデルの特定化の問題や主観的評価のバイアスによると思われる、一部直観にあわない符号を持つパラメータが存在したが、後の選択モデルを用いた分析結果から、個人の嗜好をうまく表している潜在的な態度変数を得ることができた。

3) 2) で求められた態度変数を用いて潜在セグメントモデルを推定したところ、解釈が容易で、態度変数の意味付けに矛盾しない潜在セグメントが抽出された。また、適合度も比較的良好であった。

4) 3) と同様の潜在セグメントモデルに、誤差項の系列相関を考慮したモデルを適用した。その結果、3) と同様のセグメントが抽出され、パラメータの絶対値やt値など、系列相関に影響を受けるとされるものがそれほど変化しなかった。

以上の4点より、潜在セグメントモデルを休日の買物目的地選択のパネルデータに適用することで、解釈が容易なセグメントを抽出し、モデルの適合度も大幅に改善されその有効性が確かめられた。ただし、通常のアプリアリセグメンテーションと同様に、客観的な属性を用いたときは、セグメントを規定する変数の選び方によってはその有効性が生かされないことがあり、セグメントを規定する変数の選択に関する問題は残された。今回用いたパネルデータは、データの制約上から非常に少ないサンプルで推定を行ったため、限られた説明変数の導入しかされていない。このような小サンプル少説明変数の場合には、Omitted Variableの影響や、個人の異質性などの影響によって、誤差項が比較的卓越したモデルとなり、適合度を高めることが難しい。そのようなサンプルに対して、潜在セグメントモデルを用いて適合度を高めることができるのならば、潜在セグメントモデルは、少サンプルでのモデル構築に非常に有効な技法となりうる。

本研究で提案した手法は、個人の嗜好を直接表す潜在的な態度変数を、1章で示した意思決定構造パスダイアグラムの態度指標に選択理由データを用いて推定することで、各意思決定者の異質性を表現するモデルシステムの構築を可能にした。この手法は、同一個人から数多くのデータを集める必要性がなく、態度指標が異なる個人には異なる効用パラメータを与える事を可能にしている。またこのモデルに系列相関モデルを統合し、パネルデータへの適用を試みた結果より、系列相関項を導入することはその推定パラメータにそれほど影響を与えず、定数項はノンパラメトリックに分布すると仮定することで十分である可能性を示した。しかし、本研究で用いた態度変数は、あくまで行動結果に対してなぜその場所を選んだのかという意味での選択理由であり、選択に際して何を重視してるか尋ねたものではない。また、今回用いたパネルデータはSCに来た人を対象としたチョイススペースサンプリングのデータであり、基本的にはSCに対して好意的な人たちの集まりであるとも考えられる。そのため、モデルの適合度は、認知的不協和解消行動などによって過大評価されているかもしれない。同様に今回提案したモデルの適合度を左右する潜在セグメントの数も、潜在的帰属度変数の数によって上限を規定されている。今回の事例研究では潜在変数の数を2つとした関係上、潜在セグメントの数を2としたが、モデル中の潜在セグメントの数はあくまで観測者の側からアプリアリに与えるものであり、今後最適なセグメント数を決定する手法または基準の提案が必要であると考えられる。また今後の方向として「交通問題に対する意識」などのより一般的な態度変数を求め、それに基づいて潜在セグメントモデルを適用することにより、政策変化に対する個人の反応の違いを表現したモデルが構築できると考えられる。

手法的な課題としては、セグメント帰属関数や線形構造方程式モデルは各WAVEでそのスケールを固定したが、経年的にそれらが変わる可能性は否定できないことや、選択モデルの説明変数は総合評価を用いているが、これらの将来値を得ることは不可能なので、将来予測の際には、観測可能な客観的要因の関数で置き換える必要がある。また、態度変数を求める線形構造方程式モデルと、選択モデル部が母集団に対し、線形構造方程式モデルは一つの正規母集団、選択モデルは異なる2つの潜在的母集団、というそれぞれ異なった母集団に対する仮定をおいているため、線形構造方程式モデルにも個人の異質性を導入する手法の提案などが課題としてあげられる。

以上、本研究は課題を多く抱えてはいるが、嗜好の異質性を考慮したパネル分析の足がかりとなる成果を残せたと考えられる。



#### 第4章 参考文献

- 内田敬 (1993) 情報提供を考慮した動的経路選択の交通行動分析に関する研究, 京都大学博士論文.
- 佐和隆光 (1970) 計量経済学の基礎, 東洋経済新報社.
- 渋谷政昭 (1965) 数値積分と数値微分, 山内二郎, 森口繁一, 一松信編, 電子計算機のための数値計算法 I, 第4章, 培風館.
- 西井和夫, 近藤勝直, 古屋秀樹, 栃木秀典 (1996) 多時点パネルのアトリションバイアスに関する基礎的考察: 甲府買物パネルデータを用いて, 土木計画学研究・講演集, No.19(2), pp.791-794.
- 兵藤哲朗 (1988) 需要主体の時間的行動変化を考慮した非集計形交通需要予測モデル, 交通と統計, No.18, 第4章, 交通統計研究所, pp.65-82.
- 山田菊子 (1991) 系列相関を持つ RP データと SP データを同時に用いる交通行動モデルの推定法, 京都大学大学院工学研究科修士論文.
- Amemiya, T. (1985) Advanced Econometrics, Harvard University Press.
- Bradley, M. (1992) A Practical Comparison of Modeling Approaches for Panel Data, Paper for the First U.S. Conference on Panels for Transportation Planning.
- Daganzo, C. F. and Sheffi, Y. (1982) Multinomial Probit with Time-Series Data: Unifying State Dependence and Serial Correlation Models, Environment and Planning A, Vol.14, pp.1377-1388.
- Dunn, R. and Wrigley, N. (1985) Beta-logistic Models of Urban Shopping Centre Choice, Geographical Analysis, Vol.17, No.2, pp.95-113.
- Heckman, J. J. and Willis, R. J. (1977) A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women, Journal of Political Economy, No.85, pp.27-58.
- Heckman, J. J. (1981) Statistical Models for Discrete Panel Data, In Structural Analysis of Discrete Data with Econometric Applications, Manski C. F. and McFadden D. (eds.), MIT Press, pp.114-175.
- Hensher, D. A. and Wrigley, N. (1986) Statistical Modeling of Discrete Choices in Discrete Time with Panel Data, Behavioural Research for Transport Policy, VNU Science Press, pp.97-116.
- Hensher, D. A. and Raimond, T. (1992) The Timing of Change: Discrete and Continuous Time Panels in Transportation, Paper for the First U.S. Conference on Panels for Transportation Planning.
- Hensher, D. A., Smith, N. C., Milthorpe, F. M. and Barnard, P. O. (1992) Dimensions of Automobile Demand: A Longitudinal Study of Automobile Ownership and Use, In Study in Regional Science and Urban Economics, Anselin, L. et al (eds.), Vol.22, North Holland, pp.52-86.
- Hsiao, C. (1986) Analysis of Panel Data, Cambridge University Press.
- Kim, S. G. and Mannering, F. L. (1992) Panel Data and Activity Duration Models: Econometric Alternatives and Applications, Paper for the First U.S. Conference on Panels for Transportation Planning.
- Kitamura, R. (1990) Panel Analysis in Transportation Planning: An Overview, Transportation, Vol.24 (A), No.6, pp.401-405.
- Kitamura, R. and Bunch, D. S. (1990) Heterogeneity and State Dependence in Household Car Ownership: A Panel Analysis Using Ordered Response Probit Models with Error Components, Transportation and Traffic

Theory, M. Koshi (ed.), pp.477-498.

Luce, D. (1959) Individual Choice Behavior, John Wiley and Sons.

Markus, G. B. (1979) Analyzing Panel Data, Series: Quantitative Applications in the Social sciences, SAGE Publications.

Morikawa, T. (1994) Correlating State Dependence and Serial Correlation in the RP/SP Combined Estimation Method, Transportation, 21, pp.153-165.

Murakami, E. and Ulberg, C. (1992) Current Status of the Puget Sound Transportation Panel, Paper for the First U.S. Conference on Panels for Transportation Planning.

Pas, E. I. and Koppelman, F. S. (1987) An Examination of the Determinants of Day-to-day Variability in Individuals' Urban Travel Behavior, Transportation, Vol.14, pp.3-20.



## 第5章 パネル消耗を考慮した選択モデル分析

### 5-1 本章の背景と目的

4章でも述べたように、交通行動は一般的に時間に対して独立であるとはいいがたく、時間に従属関係を伴っている場合が多い。例えば、観光目的地選択の場合、それまでの経験の有無が目的地選択に影響を与えていることは明白であり、各時点で独立に目的地の選択を行っているという仮定は成立しづらい。それに対し、クロスセクショナルデータは、行動の時間依存性などの情報を得ることができないため、基本的に時間に対して独立を仮定せざるを得ないことから、誤差項の独立性と同一性を保証するには情報不足となる場合がある。そこでパネルデータを用いた交通行動分析が近年盛んに行われるようになってきたが、パネルデータはその特長の代償としてさまざまな欠点も合わせ持っている。そのうちの一つが消耗バイアス（Attrition Bias）である。一般にパネルサンプルは、転居、死亡などの物理的要因や、調査拒否などの理由によって調査回数が増加するにつれて減少していく。この時、脱落したサンプルが分析対象となる事象について非独立、つまり、脱落したサンプルの特性と分析対象となる事象が関係を持っている場合、残留したサンプルはもはや母集団を正確に代表しなくなり、偏り（バイアス）が存在している。一般にパネル分析ではこのバイアスを消耗バイアスと呼ぶ。この問題はパネル分析の創成期から指摘されており、5-2で詳しく述べるようにパネルサンプルに生じる消耗バイアスを修正した分析手法が数多く開発されている。パネルデータを用いた4章では、個人の効用パラメータの異質性を主な研究対象としたため、サンプルの消耗を分析対象から切り離し、パネルから消耗したサンプルを排除したデータだけを用いて分析を行った。しかし、後述するがパネル消耗は4章で行った個人の異質性の分析と重要な関連を持っているため、本章ではパネル消耗に焦点を当てて研究を行う。本章ではパネルデータの消耗の特徴をまとめるとともに、これまでに提案された修正手法について概観し、新たなパネル消耗によるバイアスを修正したモデルの構築を行う。本研究で提案するモデルの特色として、後に定義する前消耗を含めた修正手法を構築し、調査への非協力行動全般を含めたモデルの提案になっている。その事例研究として4章でも用いた休日における買物目的地選択行動を分析対象とし、非観測要因を含めて、どのような要因と消耗バイアスが関連しているのかを検証すると共に、提案したモデルの有効性を検証する。

パネルデータを用いて休日の買物交通の行動特性を明らかにすることは、近年その重要度を増している。そのような行動特性を把握できることは、休日の買物交通のための公共交通機関や道路の整備と、それに対応した土地利用計画の基礎を与えることを可能とし、ひいては新規ショッピングコンプレックス立地のマーケティングや、既存ショッピングコンプレックス改善のための、交通行動から見た基礎的情報の提供が可能となる。

## 5-2 パネル消耗の特徴と既存研究のレビュー

### 5-2-1 パネルサンプルの消耗

鈴木ら（1990）は、5-1で述べた消耗バイアスが、パネルデータを用いた需要分析の最大の欠点であると指摘している。パネルサンプルの減少によるバイアスは、短期的な観点からは転居などの物理的要因の他に、回答者が回答するという行為に疲れたり飽きてしまうことなどによって調査から離脱してしまう事によって起きる。これに加えて長期的な観点からは、サンプルの調査対象地域内の転出入や年齢構成の変化に伴い、調査対象母集団の属性分布が変化してしまう場合に発生し、すべては分析の対象となる母集団と分析に用いているサンプルの属性分布が異なってしまうことに起因する問題である。さらに国内での交通調査の場合、初期のパネル分析が行われた1980年代では、交通関連の調査で回答者を具体的に特定する氏名、住所などの記入を求めることはあまり一般的でなかった。そのため、パネルデータの作成は、多時点にわたるクロスセクショナル調査を行い、その中のサンプルの同一性を年齢や性別などから推測する「疑似パネル」に頼らざるを得ないという調査上の問題があった。このような場合は、回答者がアンケートに参加するのに疲れて、調査への参加を拒否することでパネルから脱落してしまうことと、回答者が転出等の社会的事由によってパネルから脱落してしまうことの違いを、分析者側から判別するのは困難である。このように初期の国内のパネル調査では、サンプル脱落の理由を直接的に分析者が得られなかったことが、パネル消耗の分析をますます困難にしていた。

ここで注意しなければならないのは、消耗バイアスとは、パネルに滞留しているサンプルの非観測要因と選択結果の間に何らかの相関が存在する場合に生じるということである。つまり、選択モデル推定の場合、アトリションによってサンプルの数が減少し、かつパネルに残ったサ

ンプルの外的属性や選択シェアが偏ったとしても、選択に影響を与える外生変数を説明変数として用いてモデルを推定したならば、得られるパラメータは不偏性を保ち、アトリションによるバイアスは生じないことになる。例えば、交通機関選択分析において、高年齢層ほどパネル残留傾向と公共交通機関の選択率が高く、アトリションによってサンプルが大幅に減少し、高年齢者層ばかりがサンプルに残った場合でも、公共交通機関の選択シェアが母集団と比較して高いことが、年齢を説明変数に加えることによって説明されるならば、推定されたモデルのパラメータにバイアスは生じないということである。非観測要因がサンプルの減少と選択行動の両方に影響を与えている場合は、先の例で年齢という説明変数を選択モデルに導入しなかったことと同じになり、誤差項の期待値にバイアスが生じ、結果として効用関数のパラメータにバイアスが生じてしまう。ここでパネルアトリションと交通行動の関係を調査した具体的な例を挙げると、オランダで行われた交通行動全般に対する大規模なパネル調査であるDutch Mobility Panelでは、低収入層の回答率が低いこと（Golob et al.（1986））やモビリティの高いサンプルがパネルから脱落し易いという報告（Meurs and Ridder（1992））がある。Dutch National Mobility Panelではこの他に、単身者、高齢者の世帯は特に脱落しやすく、その非ランダムな脱落現象は調査初期では特に顕著で、調査が進むにつれてパネルからの脱落はランダムになって行き、WAVE6やWAVE7まで至るとほぼ完全にランダムに脱落する傾向が見られる（vsn Wissen and Meurs（1989））。モビリティの高いサンプル脱落の理由として、Meurs and Ridder（1992）はパネル調査は負担が大きいため（Dutch National Mobility Panelは1 WAVEごとに7日間の調査をする）、外出回数の多いサンプルは脱落するか、外出数を過小に答える傾向があることを指摘している。このようにモビリティの高いサンプルがパネルから多く脱落すれば、集計的な分析においてはサンプル中のトリップ数の平均値が小さくなることで、トリップ生成数を過小評価することになる。また、このサンプルを用いて交通発生をモデル分析した場合に、観測可能な要因でトリップ発生回数が説明しきれないときは、非観測要因が消耗に影響を与えていることになり、モデルの推定パラメータにバイアスが生じる可能性がある。このことに注意して考えると、ある特定の選好構造を持ち特定の選択肢の選択率が高いサンプルが、パネル調査に残留する確率が他のサンプルに比べて高いときに、母集団での選択率とパネルサンプルでの選択率との間に生じる乖離は、WAVE（パネル分析では、各調査時点のことをWAVEと呼ぶ）を重ねる毎に大きくなっていく。このようなサンプルの選択構造をモデル化できていない場合、つまり

消耗と選択双方に関連する要因を観測できない場合には、選択肢の選択シェアが母集団と異なり一種の選択肢別標本抽出のサンプルになっていると考えられ、このサンプルをモデル構築に用いた場合には、その推定値、特に選択肢固有定数項にバイアスが生じることはよく知られている（Ben-Akiva and Lerman（1985））。このようなバイアスを防ぐために、サンプルアトリションの発生を最小限にとどめるよう調査を行うために、様々な実践的方法論が試みられている。例えば調査結果を随時被験者に知らせることや、あらかじめ電話で調査への協力を確認してからサンプリングを行ったり、WAVE間に電話などで被験者にコンタクトをとるといった方法である（例えば、Murakami and Ulberg（1992）、Kumar and Replogle（1992））。これとは別に長期的なサンプルのバイアスを防ぐために、サンプルの一部を定期的に入れ替えて行き最終的には全量の入れ替えを行うPanel Rotationや、減少したサンプルに新規の追加サンプルを加えるMixed Panelなどの手法（Hensher（1987））も試みられている。一方で、消耗がおきたサンプルを用いたうえで、バイアスのない推定量を求める、統計的なモデリング手法に関する研究も数多く見られる。これらの消耗－行動分析モデルの多くは、各WAVEで消耗していった後のサンプルの属性分布を、母集団の属性分布に一致させるように重み付けを行うことで消耗バイアスを除去しようとしているものや、消耗行動と選択行動を説明するモデルの誤差相関をモデル化するものに大きく分けられる。その幾つかの例を次節で挙げる。

## 5-2-2 パネル消耗に関する既存研究のレビュー

近年のパネルデータを用いた既存研究は、パネルデータの特長を生かした上で、4-2で述べたパネルデータの欠点を修正するものが多く、その他にはデータ採集方法および採取されたデータに存在するバイアスの確認などが多い。その中で、パネルデータ独自のモデルを用いるものはパネル消耗に関するものの他には、4-2で述べたLeft-censoringの問題や活動期間の持続などのRight-censoringのものがある。特に研究が盛んに行われているのパネル消耗バイアスの問題は、見た目にも分かり易いことや、問題点が明確で離散・連続モデル構築の際の選択性バイアスの問題や層別抽出サンプルでのモデル推定の問題と類似しているため、Heckmanの修正項の導入（佐野（1990））や層別抽出サンプルによるモデリング手法（Manski and Lerman（1977））等の既存の手法の応用によって対処できる場合があることを主な理由として、これまで比較的研究が盛んであったと思われる。その消耗バイアス修正の研究を先に述べたの2系統にまとめ、

具体的な研究内容を以下に詳しく述べる。

### 1) 選択性バイアス修正法

パネル残留行動と選択行動の間に非観測要因による相関があった場合、残留したサンプルで構成されたパネルデータによって推定される選択モデルは、真の誤差項の期待値がゼロにならない危険性を含んでいる（杉恵ら1993）。そこでこのゼロからのずれをモデル化し、誤差項の期待値がゼロになるように修正しない場合、推定したパラメータにはバイアスが生じる。Heckman（1976）は離散－連続モデルを用いる際に、これと似た現象である、特定の離散量を選択したサンプルのみによって需要モデルを推定しなければならないという選択性バイアスの修正に、離散型選択モデルと需要モデルの誤差項に相関を持たせるモデルを提案し、選択性バイアス修正項を示している。このより一般的なモデルはPendyala and Kitamura（1995）に示され、誤差項の相関や状態依存性の有無などによりその推定法などがまとめられている。Heckmanの提案した手法の国内での適用例として、杉恵ら（1993）は、効用関数の誤差項に消耗バイアスの修正項を取り入れたモデルを推定している。以下杉恵ら（1993）に沿って、選択性バイアス修正法に基づいた、パネル消耗を考慮した行動モデル推定法を概説する。

$t$  時点の滞留者  $n$  の  $t-1$  時点での滞留及び離脱に関する潜在変数  $A_{n,p}^{t-1}$  及び  $A_{n,np}^{t-1}$  を、その確定項と誤差項の和で示すと、（5.1）式及び（5.2）式のようになる。

$$A_{n,p}^{t-1} = V_{n,p}^{t-1} + \zeta_{n,p}^{t-1} \quad (5.1)$$

$$A_{n,np}^{t-1} = V_{n,np}^{t-1} + \zeta_{n,np}^{t-1} \quad (5.2)$$

（5.1）式及び（5.2）式の左辺が潜在変数、右辺第1項が確定項、第2項が誤差項である。この誤差項にガンベル分布を仮定し消耗モデルを2項ロジットモデルで定式化すると、個人  $n$  の滞留確率は以下のようになる。

$$P_{n,p}^{t-1} = \frac{1}{1 + \exp\left\{-\left(V_{n,p}^{t-1} - V_{n,np}^{t-1}\right)\right\}} \quad (5.3)$$

この滞留確率を用いて、 $t$  時点での滞留者  $n$  の選択肢  $i$  に対する効用関数は、 $t$  時点の確定項、 $t-1$  時点で調査に参加しているという条件のもとでの誤差項の期待値と、平均ゼロの真の誤差項の3つの項で表される。



$$U_{in}^t = V_{in}^t + E(\epsilon_{in}^t | A_n^{t-1} > 0) + v_{in}^t \quad (5.4)$$

ここで、条件付き期待値  $E(\epsilon_{in}^t | A_n^{t-1} > 0)$  は、技術的に  $\zeta_{n,p}^{*t-1}, \zeta_{n,np}^{*t-1}$  の条件付き期待値をそれぞれ求めた上で求められることが知られている（佐野（1990））。

$$E(\epsilon_{in}^t | A_n^{t-1} > 0) = E(\epsilon_{in}^t | \zeta_{n,p}^{*t-1}, \zeta_{n,np}^{*t-1}) \quad (5.5)$$

ただし、

$$\zeta_{n,p}^{*t-1} = E(\zeta_{n,p}^{t-1} | A_n^{t-1} > 0) \quad (5.6)$$

$$\zeta_{n,np}^{*t-1} = E(\zeta_{n,np}^{t-1} | A_n^{t-1} > 0) \quad (5.7)$$

(5.5) 式を 3 変量正規分布の性質を利用して近似的に求めた式は以下のようになる。

$$U_{in}^t = V_{in}^t + \gamma_{ip} \left( \frac{P_{n,np}^{t-1}}{1 - P_{n,np}^{t-1}} \ln P_{n,np}^{t-1} + \ln P_{n,p}^{t-1} \right) + v_{in}^t \quad (5.8)$$

ただし、右辺第 2 項は消耗バイアス修正項、 $\gamma_{ip}$  は選択肢  $i$  に対する効用関数の消耗バイアス修正項のパラメータ、右辺第 3 項は誤差項である。この研究は現在 Mass Point 手法を導入し、回答バイアスまでも考慮できる形に発展している（藤原ら（1996））。ここで注意しなければならないのは、この Heckman の修正項は効用関数によって説明される変数が連続型の場合は一致性を保つが、離散選択の場合には一致性を持たない。Pendyala and Kitamura（1995）では、状態依存と選択性バイアスを考慮したより一般的に手法を提案し、Choice-based のパネルサンプルに適用して事例研究を行っている。その具体的な定式化は、個人の選択行動を規定する潜在変数と、個人の消耗に関する潜在変数を（5.9）式～（5.12）式のように表している。

$$C^* = \lambda' X + \xi \quad (5.9)$$

$$c = \begin{cases} 1 & \text{if } C^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.10)$$

$$A^* = \delta' Z + \eta c + \zeta \quad (5.11)$$

$$a = \begin{cases} 1 & \text{if } A^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.12)$$

$C^*$ ：選択行動を規定する潜在変数（効用）

$C$ ：最初の選択結果

$A^*$ ：パネル消耗行動を規定する潜在変数

$a$ ：パネル残留結果

$\lambda, \delta$ ：パラメータベクトル

$\eta$ ：選択結果とアトリションに関連度を示すパラメータ

$Z, X$ ：選択とアトリションに影響する外生変数

$\zeta, \xi$ ：誤差項

このときに次の 2 つの条件

(a)  $C^*$  は  $Z$  に対して独立でかつ  $\eta = 0$

(b) 2 つの誤差項  $\zeta, \xi$  は独立である

が、どちらか片方だけ満たされた場合、両方とも満たされた場合もしくは両方とも満たされない場合、どのような仮定のもとでのモデル化なのか、またその場合の特性について詳しく述べている。そのうえで、Choice-based サンプルの母集団代表性を考慮した選択モデルを推定している。

これらの研究の他にも、パネルデータの消耗バイアスを、選択性バイアス法を用いて修正した研究に Hausman and Wise（1979）が Gary Income Maintenance データ（Winner（1983））を用いて所得の変化の分析モデルを推定した研究や、Winner（1983）の自動車の購入意向分析を行った例が挙げられる。このほかに交通分野での研究として Kitamura and Bovy（1987）のトリップの発生回数を離散一連続モデルを用いて分析した研究や、Hensher et. al.（1992）の自動車の走行量を分析などが挙げられる。

## 2) サンプリングウェイトを考慮する手法

Manski and Lerman（1977）は、各離散選択肢の母集団における選択シェアが既知であれば、Choise-based サンプルを用いて一致性のあるパラメータ推定量を得る手法を示している。属性  $x_i$  をもつサンプルのパラメータ  $\theta$  が与えられたときの対数尤度関数を、 $\sum_i L_i(\theta, x_i)$  とすると Manski and Lerman の提案する WESML（Weighted Exogenous Sample Maximum Likelihood Estimator）推定量は、 $w_i$  を既知の正の値である重みとしたとき  $\sum_i w_i L_i(\theta, x_i)$  の最大化によって得られる。WESML 推定量は、一致性、漸近的正規性は持つが、有効性は持たない推定量である。サンプ

ルウェイトが既知であればWESML推定量はどのようなサンプリング設計にも用いることができる。この手法は通常の最尤推定法の尤度関数にウェイトをかけるだけであるので、既存のパッケージなどで容易に推定できるということが大きな利点である。この考え方を応用したパネル消耗を考慮したモデルとしては、西井ら（1995-2）が、パネルの次回調査への参加－不参加モデルを用いて、次回のパネルへの参加確率を推定し、その逆数をサンプルウェイトとして使用するモデルを提案している。しかし、パネル調査から脱落してしまったサンプルについての情報は、脱落する一時点前のものまでしかない。よって、性別などの時間に依存しない属性に関しては問題はないが、時間に依存する属性、例えば収入や職業などはパネルから脱落した時点での情報を得ることはできないため、これらの時間依存属性の一時点前の情報を用いて消費重みに使うことは、正確なサンプリングウェイトになっていないという問題がある。Brownstone and Chu（1992）では、多重内挿法（Multiple Imputations Technique）をWESML推定量と共に用いて、先に述べたサンプルウェイトが不正確になってしまうこと補正している。この手法はBrownstone and Golob（1992）でサンプルウェイトの不確実性を考慮するために用いられたもので、サンプルウェイトを特定の確率分布を持つ確率変数として考え、乱数によってサンプルウェイトをシミュレーションで与え、そのウェイトを用いてパラメータを推定することを繰り返し、その結果として一致性のある推定量を求めるという手法がとられている。

5－2－3 調査非協力層の存在とサンプルバイアス

核家族化の進展や女性の社会進出といった社会状況の変化による家庭の不在率の上昇や、調査員の確保難、人件費の高騰、調査員の安全の確保が困難になっているといった要因により、大都市を中心に実査環境が悪化してきたため、近年、調査の規模にかかわらず、郵送による調査票の回収が、これまで訪問によって行われてきたデータ収集にとって変わりつつある。また、経路や対象目的地での選択肢別抽出では、以前から現場配布郵送回収の形をとる調査がほとんどであった。調査形態別の特徴をまとめた山形（1995）によると、郵送回収は回収時のチェックができないため未記入事項が多いことが多く、精度の高いデータ収集は困難であるとされている。また、郵送回収は一般に訪問回収に比べて回収率が低く、配布形態が現地配布であるか郵送配布であるかにかかわらず、その回収率は10％～30％程度であることが多い。山形（1995）が調査方法別の特色をまとめた表の一部を表5－1に掲載する。

表5－1 主な調査方法における調査特性（抜粋）

| 方法    | 特性 | 対象者のランダム性 | 有効回収における偏り | 記入内容の質 | 記入内容の確認 | 期待しうる有効回収率 | 費用 |
|-------|----|-----------|------------|--------|---------|------------|----|
| 訪問面接  |    | 良         | 少          | 良      | 有       | 80～90％     | 極高 |
| 訪問留置  |    | 良         | 少          | 中      | 有       | 70～90％     | 高  |
| 郵送    |    | 良         | 多          | 否      | 無       | 10～30％     | 低  |
| 電話    |    | ほぼ良       | 少          | 良      | 有       | 70～90％     | 高  |
| 現場面接  |    | 否         | 多          | 良      | 有       | 低          | 低  |
| 現場調査票 |    | ほぼ良       | 少          | 中      | 無       | 60～90％     | 中  |
| 現場郵送  |    | 否         | 多          | 否      | 無       | 10～20％     | 低  |

このように郵送回収はその回収率が低い上に、その無回答者（回答者）の中に何らかの偏りがあることが指摘されている。その具体的な例として伊東（1984）は、全国幹線旅客流動調査

の予備調査で、旅行回数の多い層で高い回収率を示し、これを単純に集計すると極めて高い生成原単位となることを示した。森川ら（1993）は郊外部と都市部の訪問調査と郵送調査の比較を行い、訪問調査・郵送調査それぞれ比較的回収率の高い郊外部では郵送調査では自由トリップや短トリップの抜け落ちに伴う外出率の低下が観測される一方、回収率が訪問・郵送ともに低い都市部においては訪問調査の優位性がすでに失われている可能性を指摘している。

本研究で取り扱うパネル調査は、何度も繰り返し調査を行うため、その調査費用と労力の低減を目的として、WAVE1から郵送回収になることが多い。そのため、第1回の調査時に、先に述べたような郵送調査の特徴によって、サンプルが偏りやすく、その後パネル消耗によるサンプルの減少を考えると、WAVE1での無回答バイアスの問題は重要である。このことを考えて、アメリカなどの欧米での調査では、あらかじめ電話による参加意志の確認を行ってからパネル調査を開始する例があり（例えばMurakami and Ulberg（1992））、このような場合は有効回収率が高くなる。同様な試みをした国内の例としては、大阪湾岸パネル調査（瀬戸ら（1994））で、事前予備調査に反応したサンプルだけを対象にパネル調査を行った事例があるが、電話による調査参加への意思確認を行った例はない。ここで国内でのいくつかのパネル調査でのWAVE1での（有効）回収率を示す。



表 5－2 国内の代表的なパネル調査における初期回収状況

|                           | 配布数  | 回収数<br>(回収率)   | 有効回収数<br>(有効回収率) | 備考                 |
|---------------------------|------|----------------|------------------|--------------------|
| 甲府買物パネル<br>(WAVE 1 来訪)    | 1500 | 653<br>(53.5%) |                  | 現場配布郵送回収<br>(質問数少) |
| 甲府買物パネル<br>(WAVE 3 来訪)    | 2000 | 357<br>(17.9%) | 161<br>(8.1%)    | 現場配布郵送回収<br>(質問数多) |
| 堺所要時間パネル<br>(大浜WAVE 1 来訪) | 2290 |                | 634<br>(27.7%)   | 現場配布郵送回収           |
| 茨木駐車場パネル                  | 2757 | 856<br>(31.0%) | 831<br>(30.1%)   | 現場配布郵送回収           |

表 5－2に見られるように郵送回収による調査の（有効）回収率は、質問数の多少などにも影響されるが、10%～50%である。このように、郵送による調査票回収では調査にまったく協力してくれない層は非常に大きく、それらがランダムに発生しているという保証はない。先に挙げた森川ら（1993）では、非協力層の外出率への影響については補助調査を簡単なハガキによる調査や電話・訪問調査で行うことで解決できる可能性もあると指摘しているが、このように大きな割合を持つ非協力層の存在について、非集計分析を前提とした比較的小規模な調査では、費用の問題などからそれほど重要視されず、初期サンプルに母集団代表性があるとの仮定のもとで分析されることが多かったが、1回目の調査での無回答はサンプルの減少率においてはパネル消耗よりも高く、WAVE 1 の時点でサンプルの母集団代表性が損なわれている可能性は高いと言える。

5－3 消耗効果を考慮した選択モデル分析

5－3－1 パネル調査における消耗の概要

表 5－3 はこれまでに行われてきたいくつかのパネル調査の、消耗状況を中心に表したものである。これらのうち、調査対象・目的としてDutch National Mobility Panel（Van Wissen and Meurs（1989））やPuget Sound Transportation Panel（Murakami and Watterson（1990））はライフサイク

ルや所得の変化等の外部要因の変化が、交通行動にどのような影響を与えるかといった交通行動の動的特性全般を、South Yorkshire Panel（Goodwin（1986））は車保有とバス利用の長期的関連性を、Honolulu Transportation Demonstration Panel（Golob（1990））は時差出勤制試行の評価を、Southern California Commuter Panel（Uhlener and Kim（1992））は通勤交通行動を、Cardiff Consumer Panel（Wingley et. al（1985））は買物交通行動を、千葉モノレールパネル（鈴木ら（1990）、新谷ら（1990））、広島新交通パネル（藤原、杉恵（1992）、杉恵ら（1992））は新規交通機関の導入に伴う交通行動の変化を、堺所要時間パネル（Iida et. al（1992）、内田ら（1992））、は所要時間表示によるドライバーの学習効果をそれぞれ主な目的とした調査である。表 5－3 を

表 5－3 代表的なパネル調査のWAVE間の消耗状況

|   | 配布数          | wave1<br>回収数<br>調査年                  | wave2<br>回収数<br>調査年                | wave3<br>回収数<br>調査年 | wave4<br>回収数<br>調査年 | wave5<br>回収数<br>調査年 | wave6<br>回収数<br>調査年 |
|---|--------------|--------------------------------------|------------------------------------|---------------------|---------------------|---------------------|---------------------|
| Dutch National<br>Mobility Panel                  |              | 1764<br>1984                         |                                    | 1030<br>1986        |                     | 854<br>1988         |                     |
| Puget Sound Panel                                 |              | 1713<br>1989                         | 1391<br>1990                       | 801<br>1992         |                     |                     |                     |
| South California<br>Transportation Panel          | 2953         | 1448<br>1990                         | 833<br>1990                        | 535<br>1990         | 344<br>1990/91      | 274<br>1991         |                     |
| South Yorkshire<br>Panel                          |              | 3128<br>(世帯)<br>8252<br>(個人)<br>1981 | 1348<br><br>3045<br><br>1984       |                     |                     |                     |                     |
| Honolulu<br>Transportation<br>Demonstration Panel |              | 2297<br>1988                         | 2005<br>1988                       |                     |                     |                     |                     |
| Cardiff Consumer<br>Panel                         |              | 2012<br>1982                         | 455<br>1982                        |                     |                     |                     |                     |
| Metro Manila Panel                                | 1260<br>(世帯) | 892<br>(世帯)<br>1437<br>(個人)<br>1987  | 432<br>(世帯)<br>204<br>(個人)<br>1990 |                     |                     |                     |                     |
| 堺所要時間パネル<br>(大浜 1)                                | 2290         | 634<br>1991                          | 388<br>1991                        | 295<br>1991         | 251<br>1992         | 224<br>1992         | 137<br>1993         |
| 甲府買物パネル<br>(WAVE 1 配布)                            | 1500         | 653<br>1989                          | 221<br>1990                        | 153<br>1991         | 117<br>1992         | 86<br>1993          | 71<br>1994          |
| 広島SPパネル<br>(WAVE 1 配布)                            |              | 509<br>1987                          | 75<br>1988                         | 52<br>1990          | 37<br>1993          |                     |                     |

見るとパネル調査の消耗はその調査形態、インセンティブの有無、調査間隔、調査項目の数、



調査対象などにより一定ではない。これらの消耗に対して消耗そのものを減らすための具体的分析として、5-2-1で述べたようにインセンティブの供与、被験者とのコンタクト、調査方法の変更などによる消耗率の変化が検証されている。前述のPuget Sound Transportation Panelでは金銭的なインセンティブを増加しても、返答率が增加するという保証はないと報告されている（Murakami and Watterson（1990））。Puget Sound Transportation Panelではその他にも被験者と常にコンタクトをとるためにWAVE間でGreeting Cardなどを送っている（Murakami and Ulberg（1992））。Montgomery County Trip Purpose Panelでは無料乗車券など金銭でないインセンティブを与えたが、回収の時間的制約などによりそれほどの効果を上げていない（Kumar and Replogle（1992））。

5-3-2 消耗バイアスの定性的・定量的分析

本節では、4章でも用いた甲府パネル調査この甲府パネル調査の詳細な分析は、調査の実施主体である山梨大学の西井研究室において既に行われている。その分析の中にはもちろん消耗による影響の分析も含まれている。本節ではその中で、西井ら（1995-1）の分析を参考に消耗の影響を見る。

甲府パネル買物調査は現時点でWAVE7までの集計が終わっている。そのなかでWAVE6までのパネル調査残留状況は以下の図5-1のように表されている。

西井ら（1995-1）はWAVE3の来訪者を分析対象にして、参加層と不参加層の年齢層別構成比を例に挙げ、その違いを図示して低年齢層の不参加率の高さを指摘している。また、パネル調査への参加・不参加モデルを構築し、年齢、家族人数、ライフサイクルステージ、などといった個人・世帯属性や価値観の違いによる調査参加傾向の違いを表す変数が有意なパラメータを持っていることを示している。さらに、活動目的、購買品目、訪問箇所数などの買物場所への利便性・愛着度もパネル残留を規定する要因として有効であることを示している。これと同じデータを用いて、WAVE3に来訪しWAVE5まで残留したサンプルを例にして、その性別構成を図5-2に示す。図から分かるようにWAVE3では37.6%あった男性の比率がWAVE4で35.6%に低下し、WAVE5では31.7%まで低下していることから、サンプルの減少率は男性の方が女性よりも高いと思われる。つまりパネル調査への残留という事象は、性別という属性に対してラ

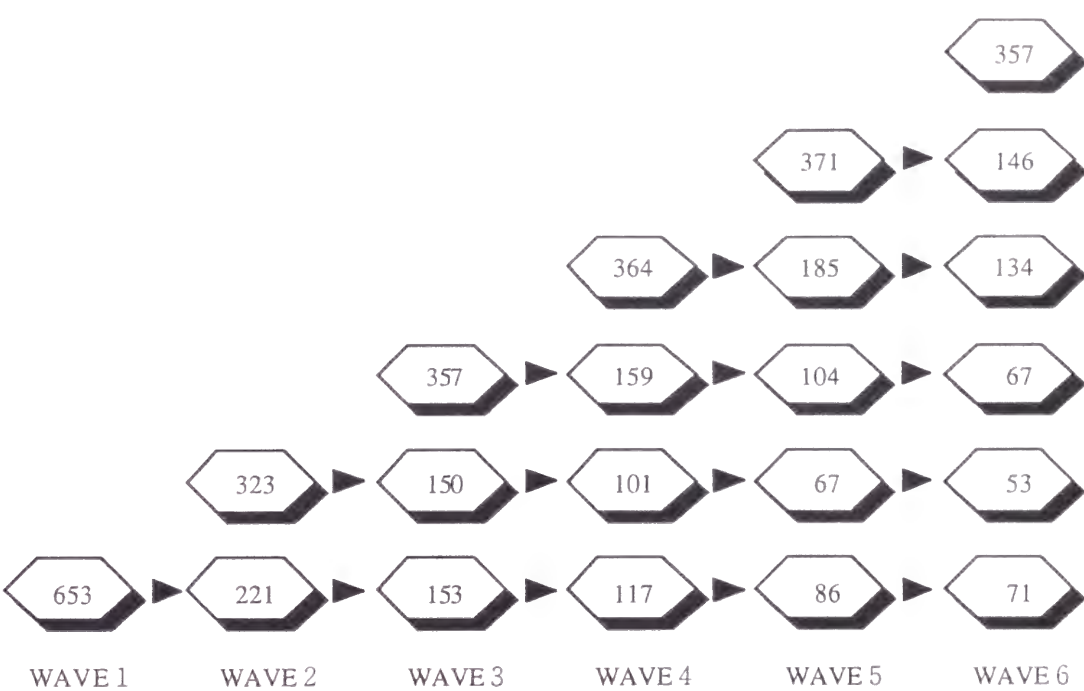


図5-1 甲府パネルでの各WAVEごとのサンプル残留状況

ンダムに生起していないと考えられる。また調査項目の中の「ここ1～2カ月最もよく利用する買物場所」のシェアを、性別ごとと全サンプルとでそれぞれWAVEごとに表5-4～6にまとめた。この集計結果では、女性と男性のそれぞれの買物場所の選択シェアに差が存在すること

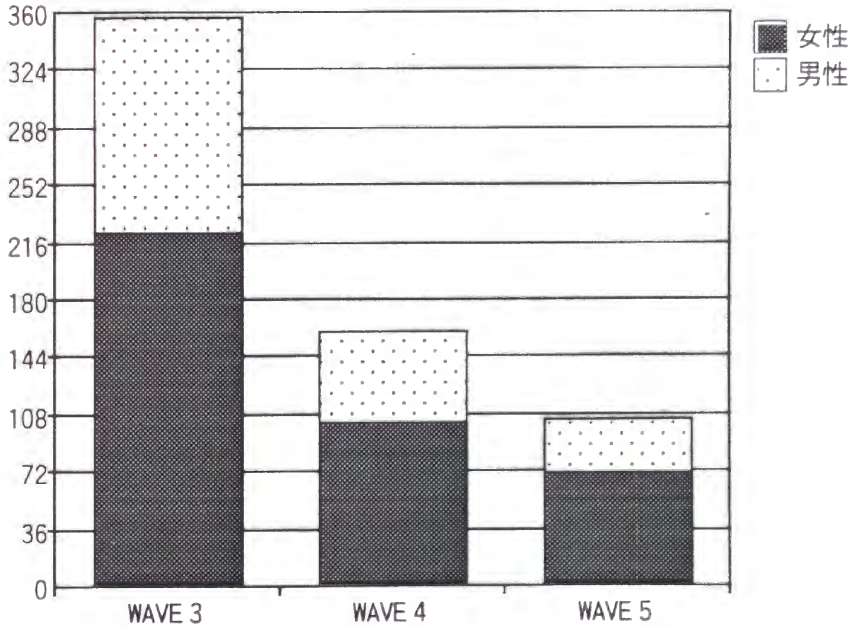


図5-2 各WAVEでのパネルサンプルの性別構成

表 5-4 (WAVE3) グループ別および全サンプルでの選択シェア

| WAVE3   | 女性    | 男性    | 全サンプル |
|---------|-------|-------|-------|
| SC      | 63.5% | 64.9% | 64.0% |
| 中心部商店街  | 8.2%  | 3.8%  | 6.5%  |
| 最寄りスーパー | 28.3% | 31.3% | 29.5% |
| 計       | 100%  | 100%  | 100%  |

表 5-5 (WAVE4) グループ別および全サンプルでの選択シェア

| WAVE4   | 女性    | 男性    | 全サンプル |
|---------|-------|-------|-------|
| SC      | 55.3% | 54.4% | 55.0% |
| 中心部商店街  | 14.6% | 10.5% | 13.1% |
| 最寄りスーパー | 30.1% | 35.1% | 31.9% |
| 計       | 100%  | 100%  | 100%  |

表 5-6 (WAVE5) グループ別および全サンプルでの選択シェア

| WAVE5   | 女性    | 男性    | 全サンプル |
|---------|-------|-------|-------|
| SC      | 49.3% | 69.7% | 55.8% |
| 中心部商店街  | 7.0%  | 6.1%  | 6.7%  |
| 最寄りスーパー | 43.7% | 24.2% | 37.5% |
| 計       | 100%  | 100%  | 100%  |

が認められる。これらの結果から、性別は買物目的地の選択とパネル残留それぞれに何らかの関連があると考えられ、このパネルデータを用いて買物場所の選択行動を分析する場合には、パネルアトリションの影響によって、サンプルが母集団（本章ではSCに買物に来ていた人を指す）を正確に表していないと言える。

ここで、非集計モデルを用いて、パネルに滞留したサンプルと脱落したサンプルの選択行動に何らかの差があるのかを確認する。選択肢は先に挙げたSC、中心部商店街、最寄りスーパーの3項で、買物目的地選択を以下に示す選択肢別の変数とSCと中心街に定数項を導入したモデルとした。

1) 自宅から買物場所への距離

2) 各買物目的地に対する回答者の交通利便性満足度指標（10段階評価）

3) 各買物目的地に対する回答者の買物利便性満足度指標（10段階評価）

ただし、全データの中から、推定に使用する各買物場所への自宅からの距離及び、買物場所に対するサンプルの買物利便性満足度及び交通利便性満足度の指標が揃っているサンプルを抽出し、WAVE3とWAVE4のデータを用いて2度以上アンケートに回答しているサンプルと1度しか調査に参加していないサンプルに分割して、それぞれについてモデルを推定し、それらの間にパラメータの相異があるかを検証する。

WAVE3の各選択肢の効用関数は以下の(5.13)式から(5.15)式のようになり、WAVE4についても同様に定義されている。

$$U_{n,SC}^{wave3} = \beta Z_{SC}^{wave3} + \varepsilon_{n,SC}^{wave3} \quad (5.13)$$

$$U_{n,C}^{wave3} = \beta Z_C^{wave3} + \varepsilon_{n,C}^{wave3} \quad (5.14)$$

$$U_{n,SP}^{wave3} = \beta Z_{SP}^{wave3} + \varepsilon_{n,SP}^{wave3} \quad (5.15)$$

ただし、

$U_{n,SC}^{wave3}$ ,  $U_{n,C}^{wave3}$ ,  $U_{n,SP}^{wave3}$ ; それぞれSC, 甲府市中心街, 最寄りスーパーのWAVE3の効用

$Z_{SC}^{wave3}$ ,  $Z_C^{wave3}$ ,  $Z_{SP}^{wave3}$ ; WAVE3での各選択肢の選択肢別変数

$\beta$ ; 未知パラメータベクトル

これらの効用関数の誤差項にIIDガンベル分布を仮定し、多項ロジット型の選択確率を定式化し、それを尤度関数とする最尤推定法により未知パラメータを求めた。以下の表5-7は調査への参加回数が1度のサンプル(WAVE3ないしWAVE4において来訪者として初めて調査に参加し、次のWAVEで脱落しているサンプル)集合についての推定結果、表5-8はWAVE3と4の調査に両方とも参加したサンプルを用いた推定結果、表5-9は調査への参加回数は考慮せずに、WAVE3ないしWAVE4の調査に参加しているサンプル全てをプールした場合の推定結果である。これらの推定では総ての観測された個人を独立とみなして推定をしている。また、スケールパラメータは3章でSP・RPモデルに適用したときと同様に、WAVE3とWAVE4の誤差項の分散の大きさの違いを調整するパラメータであり、以下の関係が成立している。

表 5－7 調査参加回数が一度のサンプルによる推定結果

| 変数名            | 推定値※ 1         |
|----------------|----------------|
| S C 定数項        | 0.309 (1.7)    |
| 中心街定数項         | -1.15 (-3.0)   |
| 距離             | 0.0083 (0.8)   |
| 交通利便性満足度       | 0.113 (2.3)    |
| 買物利便性満足度       | 0.117 (1.5)    |
| スケールパラメータ      | 1.77 (1.4) ※ 2 |
| サンプル数          | 153            |
| $\bar{\rho}^2$ | 0.250          |

※ 1 ( ) 内は t 値 ※ 2 1 を基準とした t 値

表 5－8 WAVE 3 と 4 の両方に参加したサンプルによる推定結果

| 変数名            | 推定値※ 1          |
|----------------|-----------------|
| S C 定数項        | 0.0877 (0.3)    |
| 中心街定数項         | -1.13 (-0.1)    |
| 距離             | -0.0313 (-1.1)  |
| 交通利便性満足度       | 0.426 (2.0)     |
| 買物利便性満足度       | 0.259 (2.0)     |
| スケールパラメータ      | 0.837 (0.4) ※ 2 |
| サンプル数          | 131             |
| $\bar{\rho}^2$ | 0.808           |

※ 1 ( ) 内は t 値 ※ 2 1 を基準とした t 値

表 5－9 WAVE 3 と 4 に参加したサンプル総てをプールした場合の推定結果

| 変数名            | 推定値※ 1         |
|----------------|----------------|
| S C 定数項        | 0.380 (2.6)    |
| 中心街定数項         | -0.863 (-2.4)  |
| 距離             | -0.0062 (-0.7) |
| 交通利便性満足度       | 0.175 (3.7)    |
| 買物利便性満足度       | 0.0852 (1.7)   |
| スケールパラメータ      | 2.05 (1.5) ※ 2 |
| サンプル数          | 284            |
| $\bar{\rho}^2$ | 0.490          |

※ 1 ( ) 内は t 値 ※ 2 1 を基準とした t 値

$$\text{var}(\varepsilon^{\text{wave3}}) = \mu^2 \text{var}(\varepsilon^{\text{wave4}}) \quad (5.16)$$

この 3 つの推定結果を比較すると、調査に 2 度とも参加したサンプルは定数項がいずれも有意でないが、サンプルをプールした場合は両方の定数項が有意になる。調査に一度しか参加しなかったサンプルは距離のパラメータが正の値になったが、2 度とも参加したサンプルは有意ではないが負となった。交通・買物利便性のパラメータは調査に一度しか参加しないサンプルではほぼ同等の値を持っているにも関わらず、2 度とも参加したサンプルは交通利便性の値が卓越していることに気づく。スケールパラメータの推定値は、1 の場合 WAVE 3 と 4 のスケール（誤差項の分散）が同一であることを示しているが、調査に一度しか参加しないサンプルでは 1 より小さく、2 度とも参加したサンプルでは 1 よりも大きくなり、逆の結果を示している。このように様々なパラメータでその違いが露見し、各パラメータについて t 検定を行った結果でもほとんどのパラメータが同一性を棄却され、サンプルを分離したモデルとサンプルをプールしたモデルの尤度比検定によっても、5 % の危険率で両サブグループの同一性は棄却された。よってパネル調査に連続して参加するサンプルと、一度きりしか参加しなかったサンプルではその選択構造が異なると思われる。



### 5-3-3 確率的消耗モデルを用いたパネル分析

ここでは、パネル滞留をモデル化し、パネル消耗によるサンプルのバイアスを考慮した選択モデルの構築を行う。提案するモデルの基本的なコンセプトは、これまでのパネルアトリションの研究で報告されている「パネルからの脱落は確率的に生じ、かつ個人の属性や参加したWAVEの数によって変動する」という結果を基に、各WAVE間の消耗確率がそれぞれ変動し、かつ個人属性をモデルの説明変数として取り込めるモデルを提案し、消耗現象全体を包括的に扱うことが可能なモデルシステムの構築を目標とする。具体的には機械などの故障発生確率をモデル化する際に良く用いられる生存時間モデルをベースにモデルを構築する。一般に生存時間モデルは、以下の3種類の仮定のいずれかに基づいている（脇本（1984））。

- 1) 製品が使いはじめてから時間が経過するにつれて故障が起こりやすくなる。
- 2) 製品の故障の起こりやすさは使い初めてからの経過時間に無関係である。
- 3) 初期的な故障が出尽くして、時間が経過するにつれて製品の故障が起こりにくくなる。

一般にワイブル分布などを用いて分析される生存時間モデルは、上の1)の仮定に基づくものである。3)の初期的な不良による製品の故障は、時間が経過するにつれて初期的な不良が出尽くして製品の故障が起こりにくくなるもので、パネル消耗が最初のWAVE間では特に消耗率が高く、徐々に消耗率が減少して最終的にはほぼ一定の人数に落ちつくという現象と類似している。本研究ではそのような過去の研究の報告より、初期的な不良による製品の故障をパネル消耗と類似の現象であると考え、パネル消耗分析に適用するものである。よって以下で初期不良モデルを生存時間モデルに基づいて導出し、パネル消耗確率を与えるモデルとし、消耗確率より求められた母集団でのサンプルウェイトを用いた、個人の誤差項の系列相関を認める選択モデルと同時に推定する手順を示す。

#### (1) パネル消耗モデル

5-2で概観した既存の研究によると、パネルからの離脱確率は、個人がそれまで何回調査に参加しているかによって変動し、パネル調査への参加回数が増えるにつれて離脱の割合は減ってゆき、最終的には安定した状態に到達して行くと報告されているものが多い（例えば、Wissen and Meurs（1989））。そこでパネル調査から離脱する行動を以下のように仮定する。個人が調

査票を受け取った時点からパネル調査より離脱するまでの時間は、一人一人異なるが、同一の分布に従っているとする。そこで $M$ を確率変数とし、その密度関数を $f(\tau)$ 、分布関数を $F(\tau)$ とすると、ある個人が調査票を受け取ってから $\tau$ 時間経過したという条件のもとで、それに続く微小時間内にパネル調査より脱落する確率は、 $\Delta\tau$ が十分小さいとき、次式で与えられる。

$$P(\tau < M < \tau + \Delta\tau \mid M > \tau) = \frac{F(\tau + \Delta\tau) - F(\tau)}{1 - F(\tau)} \cong \frac{f(\tau)}{1 - F(\tau)} \Delta\tau \quad (5.17)$$

ここで、

$$h(\tau) = \frac{f(\tau)}{1 - F(\tau)} \quad (5.18)$$

とおくと、 $h(\tau)$ は、 $\tau$ 時点でパネル調査に滞留しているという条件のもとで、続いておこる微小時間にパネル調査から離脱する確率密度になる。そこで、先に述べたように初期不良による故障のアナロジーに基づき、 $h(\tau)$ を時間の経過につれて調査から脱落する確率が減少する、つまり $h(\tau)$ を $\tau$ の減少関数

$$h(\tau) = a \times b^{-a\tau} \quad a, b > 0, \tau > 0 \quad (5.19)$$

ただし、

$a$ : サンプルが微小時間 $\Delta\tau$ にパネルから離脱する確率に対する時間のパラメータ

$b$ :  $h(\tau)$ の分布形を定める定数

と定義すると、 $\tau$ 時点でパネルサンプルが調査から離脱する確率の分布関数は

$$F(\tau) = 1 - \exp\left(\frac{b^{-a\tau}}{\ln b} - \frac{1}{\ln b}\right) \quad (5.20)$$

で与えられる。これはパネル調査からの離脱現象を連続時間で表す分布関数を表している。ただし、パネル調査は時間に対し離散的であるため、離散時点 $\tau-1$ までパネルに滞留し $\tau$ 時点で調査より離脱する確率は

$$P_{\tau-1, \tau}(\text{dropout}) = F(\tau) - F(\tau-1) \quad (5.21)$$

となる。既存の研究によると、個人がパネル調査に参加している期間は、個人や家庭の属性に影響を受けていること（例えばWissen and Meurs（1989））や、これまで参加したWAVE数によっても変動すること（例えば西井ら（1995-2））が、報告されている。これを、 $h(\tau)$ の分布形

を定める定数 $b$ がサンプルの個人や家庭の属性によって異なることを意味しているものと考え、個人 $n$ の属性ベクトルに影響を受ける変数 $A_n$ を、非負条件を満たすようにロジット変換して(5.20)式に代入すると以下のような形に表される。

$$F_n(t) = 1 - \exp \left[ \frac{\left\{ \frac{\exp(A_n)}{\exp(A_n) + 1} \right\}^{-a\tau}}{\ln \left\{ \frac{\exp(A_n)}{\exp(A_n) + 1} \right\}} - \frac{1}{\ln \left\{ \frac{\exp(A_n)}{\exp(A_n) + 1} \right\}} \right] \quad (5.22)$$

この式に基づいて、個人 $n$ がある時点 $t$ までパネルに滞留し、かつ次のパネル調査に不参加となる確率 $P_{n,t,t+1}^*(dropout)$ を求めると

$$P_{n,t,t+1}^*(dropout) = F_n(t+1) - F_n(t) \quad (5.23)$$

となる。

## (2) 選択モデル

ある時点 $t$ で個人 $n$ の選択肢 $i$ に対する効用を以下のように定義する。

$$U_{n,i}^t = \mathbf{B}\mathbf{Z}_{n,i} + v_{n,i}^t \quad (5.24)$$

ただし、

$U_{n,i}^t$  : 選択肢 $i$ に対する個人 $n$ の $t$ 時点での効用

$\mathbf{Z}_{n,i}$  : 選択肢 $i$ の効用に影響を与える属性ベクトル

$v_{n,i}^t$  : 個人 $n$ の選択肢 $i$ の効用に影響する測定不能な要因を表す誤差項

$\mathbf{B}$  : 未知パラメータベクトル

ここで、5-2で述べたように同一個人から時系列的に得たデータには、一般に系列相関と状態依存が存在するため、各時点間で独立で同一の誤差項を仮定するのは現実的でない。そこでMorikawa (1994)で提案されている、誤差項を加法的に分解し、同一個人に時点間で共通の誤差項と時点間で独立な誤差項に分解するモデルを導入する。具体的には時点 $t$ の効用関数を

$$U_{n,i}^t = \mathbf{B}\mathbf{Z}_{n,i} + \theta_{n,i} + \varepsilon_{n,i}^t \quad (5.25)$$

とし、時点 $t+1$ の効用関数を

$$U_{n,i}^{t+1} = \mathbf{B}\mathbf{Z}_{n,i}^{t+1} + \theta_{n,i} + \varepsilon_{n,i}^{t+1} \quad (5.26)$$

ただし、

$\theta_{n,i}$  : 各個人の時点間で共通の誤差を表す項

$\varepsilon_{n,i}^t$  : 各個人の時点間で独立な誤差を表す項

とする。

ここで、(5.22)式で用いたパネル残留傾向の違いを表す変数 $A_n$ を以下のように定義する。

$$A_n = \mathbf{\Gamma}\mathbf{X}_n \quad (5.27)$$

ただし、

$\mathbf{X}_n$  : パネル残留に影響を与える個人 $n$ の属性ベクトル

$\mathbf{\Gamma}$  : 未知パラメータ

また、先に述べたように、パネル残留行動と選択行動の間に存在すると考えられる相関を、パネル残留確率を変化させる変数 $A_n$ と各時点の選択行動を規定する効用との間に共通の攪乱項を導入することによって解決する。つまり $t$ 時点の効用と $A_n$ を、それぞれ以下のように定義し直す。

$$U_{n,i}^t = \mathbf{B}\mathbf{Z}_{n,i} + \theta_{n,i} + \lambda_n + \varepsilon_{n,i}^t \quad (5.28)$$

$$A_n = \mathbf{\Gamma}\mathbf{X}_n + \lambda_n \quad (5.29)$$

ただし、

$\lambda_n$  : 個人 $n$ の効用 $U_{n,i}^t$ とパネル残留を規定する変数 $A_n$ に共通の影響を与える攪乱項

これらより、2つの系列相関項 $\theta_{n,i}$ と $\lambda_n$ が既知のときの、時点 $t$ で個人 $n$ の選択肢 $i$ の選択確率は、 $\varepsilon_{n,i}^t$ にI.I.D.ガンベル分布を仮定することで次の式のように表すことができる。

$$P_{n,t}(i | \theta_{n,i}, \lambda_n) = \frac{\exp(\mathbf{B}\mathbf{Z}_{n,i} + \theta_{n,i} + \lambda_n)}{\sum_{k \in C_n} \exp(\mathbf{B}\mathbf{Z}_{n,k} + \theta_{n,k} + \lambda_n)} \quad (5.30)$$

ただし、 $C_n$ は個人 $n$ の選択肢集合

パネル滞留確率についても同様の定式化により

$$P_{n,t,t+1}^*(dropout) = F_n(t+1|\lambda_n) - F_n(t|\lambda_n) \quad (5.31)$$

と変形される。

### (3) 消耗バイアスを考慮した選択モデル

パネルアトリションによるパラメータのバイアスを考慮した選択モデルの構築をこれまでに定式化したモデルを用いて行う。5-2で述べたようにパネルアトリションの影響によって、パネルサンプルはWAVEが進むに従って、調査に協力的なサンプルの抽出率が徐々に大きくなった層別抽出に近似できると考えると、各WAVEでのサンプルアトリションを考慮した上での尤度は、消耗傾向の同一な個人を同一の抽出層に属すると仮定し、一般層別抽出標本による尤度と同様の考え方によって定義する。ここで一般層別抽出による尤度関数をBen-Akiva and Lerman (1987)に従って定式化する。まず、 $g$ 層に属する個人が選択結果 $i$ 、特性 $Z$ を有する尤度は(5.32)式のように $g$ の条件付き確率となる。

$$f(i, Z|g) = f(i, Z) / Q(g) \quad (5.32)$$

ただし、

$f(i, Z|g)$  :  $g$ 層において選択結果 $i$ 、属性 $Z$ を有するサンプルが得られる尤度

$f(i, Z)$  : 選択結果 $i$ 、属性 $Z$ を有するサンプルが得られる尤度

$Q(g)$  :  $g$ 層の母集団シェア

である。よって $g$ 層のサンプル内シェアを $H(g)$ とすると、あるサンプルが $g$ 層からのものであり、かつ選択結果 $i$ 、属性 $Z$ を有するという尤度 $f(i, Z)$ は、(5.33)式のようになる。

$$f(i, Z) = f(i, Z|g) H(g) = \frac{P(i|Z, \beta) P(Z) H(g)}{Q(g)} \quad (5.33)$$

ただし、

$P(i|Z, \beta)$  : 属性 $Z$ 、パラメータ $\beta$ のもとで選択肢 $i$ を選択する条件付き確率

$P(Z)$  : 属性 $Z$ の母集団分布

である。よって全サンプルについての尤度関数は(5.34)式のようになり、

$$L(\beta) = \prod_{g=1}^G \prod_{n=1}^{N_g} \prod_{i \in C_n} \frac{P(i|Z_{n,i}, \beta)^{\delta_{in}} P(Z_n) H(g)}{Q(g)} \quad (5.34)$$

ただし、

$P(i|Z_{n,i}, \beta)$  :  $\beta$ および $Z_{n,i}$ が与えられたときの選択肢 $i$ の選択確率

$\delta_{in}$  : 個人 $n$ が選択肢 $i$ を選択していた場合1、それ以外0

$H(g)$  :  $g$ 層のサンプル中でのシェア

$Q(g)$  :  $g$ 層の母集団中でのシェア

$N_g$  :  $g$ 層に属するサンプル数

$\beta$  : 未知パラメータ

$G$  : 層の総数

対数尤度関数は(5.35)式のようになる。

$$L^*(\beta) = \sum_{g=1}^G \sum_{n=1}^{N_g} \sum_{i \in C_n} \delta_{in} \ln P(i|X_n, \beta) + \sum_{g=1}^G \sum_{n=1}^{N_g} \ln P(X_n) + \sum_{g=1}^G N_g \ln H(g) - \sum_{g=1}^G N_g \ln Q(g) \quad (5.35)$$

(5.31)式において右辺第1項は、単純無作為抽出における対数尤度関数である。また、第2項及び第3項はパラメータベクトル $\beta$ を含まないので尤度関数の最大化には関係しない。第4項の母集団シェア $Q(g)$ は一般的には求めることができないが、本研究では先に定式化したパネル滞留モデルから、調査票を配った時点、すなわち本研究で言うところの母集団の人数の期待値を求めることができ、疑似母集団シェアを計算できる。このように、このモデルシステムでは生存時間モデルを用いて、最初から調査を拒否したサンプルの数を推定することができるため、通常は未知である母集団中の属性シェアを推定して、サンプルウェイトを求めることができることに特色がある。

パネル滞留モデルを用いて消耗確率が同一な層の $t$ 時点までのパネル滞留確率

$$P_{t,g}^*(stay) = 1 - F_g(t) \quad (5.36)$$

が与えられたとき、各時点でサンプル中の $g$ 層のシェアと滞留確率の間に以下のような関係が成立している。

$$\hat{N}_{0,g} = \frac{N_{t,g}}{P_{t,g}^*(stay)} \quad (5.37)$$



ただし、

$\hat{N}_{0,g}$ ：調査票配布時点（以下WAVE 0 と呼ぶ）での  $g$  層に属する人数の推計量

$N_{t,g}$ ： $t$  時点での  $g$  層の残留人数

この関係を用いると  $g$  層の WAVE 0 つまり母集団でのシェアを以下のように推定できる。

$$Q(g) = \frac{\hat{N}_{0,g}}{\sum_{j=1}^G \hat{N}_{0,j}} \quad (5.38)$$

この  $\hat{N}_{0,g}$  はどの時点の値を用いて推定するかによって変動するため、本来的には各時点ごとの値の平均値を用いることが望ましいと考えられるが、ここでは簡単のため、各時点ごとに求められる推計値を、各時点ごとの母集団シェアの推定値として適用する。このように得られた  $Q(g)$  を用いて、系列相関およびパネルアトリションを考慮した、 $g$  層に属する個人  $n$  のパネル調査への参加・不参加と全時点の目的地選択の同時確率は次のようになる。

$$P_{g,n} = \int \int \prod_{t=1}^T (L_n^{t-1} \cdot C_n^t) \cdot (P_{g,n,T,T+1}^*(dropout | \lambda_n)) \cdot f(\lambda) \cdot f(\theta) d\lambda d\theta \quad (5.39)$$

ただし、

$$L_n^t : \frac{P_{g,n,t,t+1}^*(stay | X_n, \lambda_n) P(X_n) H(g)}{Q(g)} \quad (5.40)$$

$$C_n^t : \frac{P_{g,n,t}(i | Z_n, \lambda_n, \theta_n) P(Z_n) H(g)}{Q(g)} \quad (5.41)$$

$T$ ：個人  $n$  がパネルに滞留した WAVE の最大数

$f(\lambda)$ ： $\lambda$  の確率密度関数

$f(\theta)$ ： $\theta$  の同時確率密度関数

ここで、(5.39) 式中には  $L_n^0$  が存在するが、これが仮定の WAVE 0 での人数の推定値を用いて、最初から調査に参加しなかったサンプルの割合を示している。最終的には (5.39) 式を  $f(\lambda)$ 、 $f(\theta)$  にそれぞれ標準正規確率密度、多変量独立標準正規確率密度を仮定し数値的に積分を評価することで、個人の選択確率をパラメトリックに定式化し、最尤推定法を用いて各未知パラメータを推定することになる。この推定法による得られるパラメータは、総ての仮定が

正しいとすると一致性を持つことができる。

#### 5-3-4 パネル消耗によるバイアスを修正した買物目的地選択モデル

本節では、5-3-3 で定式化した消耗・選択モデルを 5-3-2 で分析した甲府パネルデータを用いて特定化する。先に述べたように甲府パネルデータは、現在まで 7 WAVE のデータが揃っている。しかし、WAVE 1 と WAVE 2 については買物目的地選択に必要なデータが含まれていないことなどから、WAVE 3、WAVE 4、WAVE 5 のデータのみを用いてモデルを構築する。買物目的地選択のモデルは 5-3-2 と同じ 3 項選択とし、以下の同様な 3 種類の説明変数と S C と中心街の定数項を導入した。

- 1) 自宅から買物場所への距離
- 2) 各買物目的地に対する回答者の交通利便性満足度指標（10 段階評価）
- 3) 各買物目的地に対する回答者の買物利便性満足度指標（10 段階評価）

(5.27) 式に示される消耗モデルの説明変数は、パネル滞留行動に関連する属性を全て含むように設定されねばならないが、抽出層の数が、用いた説明変数の数とその属性のカテゴリ一数に応じて指数的に増えるため、今回の事例研究では、西井ら (1995-1) の研究を参考にして様々な特定化を行い比較検討をした後、5-3-2 でサンプル中の選択シェアの変動に特徴が見られた性別と定数項の線形和とした。各 WAVE で女性のサンプルはそれぞれ 100 人、48 人、33 人存在し、男性はそれぞれ 35 人、16 人、7 人存在することになる。提案したパネルアトリションモデルと系列相関を同時に考慮したモデルの推定結果を表 5-10 に示した（以下モデル 1）。それぞれ変数名中の数字は WAVE の数を表している。また比較のため、系列相関項を導入し、性別を効用関数の説明変数として加えたモデル（以下モデル 2）と、系列相関項を導入し、性別を効用関数に用いないモデル（以下モデル 3）2 つの推定結果をを表 5-11 に示した。

この 3 つのモデルの効用関数の推定値を比較すると、モデル 1（本研究で提案したモデル）は、モデル 3（性別を説明変数に用いないモデル）と似たパラメータの推定値を示している。これは女性ダミー変数を用いた消耗モデルによる重み付けが選択モデルのパラメータにそれほど影響を与えていないためと考えられる。このことは表 5-11 に示したモデル 2 の結果で女性ダ

ミーが有意なパラメータを持たないことからいえる。ただ、モデル1をモデル3と比較すると、中心街定数項が有意でなかったものがモデル1ではWAVE3とWAVE5で有意になった。同様にモデル1ではスケールパラメータが1に対して有意となり、WAVEごとの誤差分散の差異が明確に現れるようになった。モデル2と3を比較した場合、モデル2の推定結果は、サンプルの3/4が女性であることからモデル3の定数項を性別のパラメータと定数項に分解した値になっていると見ることができる。その他にモデル2と3の推定値の違いとして、モデル2の定数項がモデル3と比較して総て小さくなる方向に変化している。これは性別のパラメータが正であるために生じたことと思われる。消耗モデルについて考察すると、モデル1のパネル残留モデルの女性ダミーのパラメータの符号が正になっている。このことは、女性がパネルに残留しやすいことを示しており、先に行った集計分析で予想された結果に一致する符号である。このパネル残留モデルを取り出して各WAVEごとの残留確率を計算すると、実際の女性の残留率

表5-10 パネル消耗・目的地選択同時推定モデルの推定結果（モデル1）

| 変数名          | 推定値※1          |
|--------------|----------------|
| SC定数項3       | 1.02 (4.8)     |
| SC定数項4       | 0.630 (2.5)    |
| SC定数項5       | -0.0834 (-0.4) |
| 中心街定数項3      | -0.582 (-2.1)  |
| 中心街定数項4      | 0.349 (1.2)    |
| 中心街定数項5      | -2.21 (-2.2)   |
| 距離           | -0.577 (-6.3)  |
| 交通利便性満足度     | 0.122 (2.2)    |
| 買物利便性満足度     | 0.264 (3.6)    |
| スケールパラメータ3 4 | 1.92 (2.6) ※2  |
| スケールパラメータ3 5 | 0.658 (1.5) ※2 |
| アトリション定数     | 1.93 (8.7)     |
| 性別ダミー ※3     | 1.66 (7.0)     |
| a (定数)       | 0.540 (122.7)  |
| 観測数          | 239            |

※1：（）内はt値 ※2：性別ダミー：女性1，男性0 ※3：1を基準としたt値

がWAVE3からWAVE4にかけてが48%、WAVE4からWAVE5が68.8%であるのに対し、モデルから計算される推計値はそれぞれ57.8%、57.1%であった。男性の実際の残留率は45.7%、43.8%であるのに対し、モデルの推計値はそれぞれ54.7%、52.2%などとなり、このモデルによる予測値はサンプル全体の残留率の特徴が薄れ、平均をとったような値になっていて、あまりモデルの推計精度が高いとは言えない。ここで、本モデルの特徴である事前消耗率から計算される、事前残留率の値をみると女性で57.8%男性で56.9%残留すると推定された。これは表5-2に示した値と比べてかなり大きな値であり、消耗モデルのフィットがあまり良くない原因はこの事前消耗率による可能性が高いと思われる。これは配布数をモデル推定の際の情報として用いなかったことが原因となっていると思われるため、配布数を

表5-11 効用関数の説明変数に性別を加えたモデルと加えないモデルの推定結果（モデル2及びモデル3）

|              | (モデル2)          | (モデル3)          |
|--------------|-----------------|-----------------|
| 変数名          | 推定値（t値）※1       | 推定値※1           |
| SC定数項3       | 0.722 (1.94)    | 1.04 (5.15)     |
| SC定数項4       | 0.279 (0.69)    | 0.654 (2.13)    |
| SC定数項5       | -0.550 (-0.73)  | -0.176 (-0.28)  |
| 中心街定数項3      | -1.25 (-2.14)   | -0.584 (-1.51)  |
| 中心街定数項4      | -0.265 (-0.48)  | 0.367 (1.00)    |
| 中心街定数項5      | -3.57 (-1.84)   | -2.53 (-1.72)   |
| 性別（SC）※2     | 0.438 (1.16)    |                 |
| 性別（中心街）※2    | 0.816 (1.50)    |                 |
| 距離           | -0.571 (-4.54)  | -0.604 (-5.07)  |
| 交通利便性満足度     | 0.106 (1.99)    | 0.124 (2.24)    |
| 買物利便性満足度     | 0.263 (3.30)    | 0.286 (3.55)    |
| スケールパラメータ3 4 | 2.62 (0.87) ※3  | 1.78 (1.00) ※3  |
| スケールパラメータ3 5 | 0.533 (1.86) ※3 | 0.599 (1.58) ※3 |
| 観測数          | 239             | 239             |
| $\bar{p}^2$  | 0.255           | 0.258           |

※1：（）内はt値 ※2：性別ダミー：女性1，男性0 ※3：1を基準としたt値

外生的に与えることができる形にモデルを改良することでその適合度の改善は行えると考えられる。モデル全体としての適合度指標はアトリションを考慮したモデルの場合、消耗確率や仮想のサンプルの消耗現象を含めた尤度最大化を行っているため、通常の尤度比指標を用いることができないため、表に示してはいないが、効用関数についてはそのパラメータの推定値から、モデル3と同程度の適合度は確保していると考えられる。

#### 5-4 本章のまとめ

本研究はサンプルのアトリション現象をモデル化することで、パネルサンプルのアトリションによるバイアスを修正するモデルの提案を行った。そのとき、最初から調査への参加を拒否したサンプルの行動をモデルシステムの中に組み込み、その影響を考慮したパネルサンプルの重み付けを行った。具体的には、サンプルのパネル調査への残留傾向に影響する属性の線形和を用いてパネル消耗（滞留）モデルを構築し、それより母集団での属性シェアを定式化し、層別抽出によるサンプルの重みを考慮した尤度関数に適用して、アトリションの影響を考慮したパネル消耗（滞留）モデルと選択モデルの同時推定システムを構築した。そのとき、パネル残留に影響する潜在変数と時系列的選択の効用関数に個人特有の誤差項を導入し、2つの行動が関連した場合の影響をパラメータの推定量から除去した。

事例研究として、甲府パネル調査の買物目的地選択データを用いて、性別を消耗モデルの説明変数として採用したモデルを推定し、個人の性別を効用関数に導入したモデルや、性別を情報として導入しないモデルと比較を行った。提案したモデルは、ほとんどのパラメータが性別を説明変数に用いないモデルと同様の値を示したが、性別を説明変数に導入するだけでは起こらなかったWAVE間のスケールパラメータや中心街定数項の推定値が一定の変化をした。このことから、提案したモデルは、説明変数を追加するモデルによる推定値の変化とは異なるパラメータのバイアスの修正が行われた可能性があると言える。ただし、本研究はたかだか3回のパネル調査のデータを用いてモデルを推定したものであって、パネルの消耗（滞留）状況はこの先に異なった変化があることが予見され、パネル滞留モデルの信頼性はその説明変数の少なさとともに未だ高いとは言えない。現に配布枚数を情報として用いない今回提案したモデルでは事前消耗の確率をほとんど説明しきれていなかったため、消耗モデルの適合度は全体として大変悪いものとなった。これらを考慮した上で、事例研究を通じたより一層のパネル滞留モデル

の改善の必要性があると考えられる。また、提案モデルは個人を属性別にカテゴリー分けした上でそれぞれの消耗率を推定する構造になっているため、規定カテゴリーを増やすことで計算負荷が指数的に増加して行くことも、改善の余地を残している。

また、これとは別にパネルデータに特有に存在する各種のバイアスの問題もあるため、それらの現象把握とともに包括的なパネルデータを用いたモデルを構築することが今後の課題になってくると考えられる。例えば、近年社会統計分野で発展のある欠損値データの処理技法を援用したモデル分析をパネルアトリション分析だけでなく、不備回答などの有効サンプル減少の原因となる問題の解決に用いることも有望な展開であると考えられる。今回事例研究に使用した甲府パネル調査は調査開始から7年を数える国内でも最大規模のパネル調査であり、7年の間には変化を起こしている社会経済環境もある。甲府パネル調査はリフレッシュサンプルを精力的に追加して、社会環境の変化によるサンプルのバイアスを防いでいるが、調査費用の観点からも、そのようなサンプルバイアスの除去も長期に渡ることがあるパネル調査には必要な方法論であることは間違いない。

技術的な問題として、系列相関をモデル中に導入する際に、積分を伴うため計算負荷の増大も今回の計算を進める上での問題点としてあげることができるが、現在のめざましい計算機の実力の発展によって解消の可能性はある。ただし、パネル調査は回を重ねる毎にパネルサンプル1人当たりの情報量は増え続けることや、より包括的な動的モデルを構築するときには、より一層の計算負荷がかかると思われるため、今後実用性を視野に入れていく場合に、この問題は計算機の発達に頼ることなく解決する必要性がある。

これらのようなまだ多くの課題を抱えてはいるが、本章で提案したモデルは、パネルアトリションと無回答によるバイアスを包括的に取り扱うためのモデル研究の端緒として、意義のある結果を残すことができたと考えられる。



## 第 5 章 参考文献

Pendyala, R. M. and Kitamura, R. (1995) Weighting Methods for Attrition in Choice Based Panel, 交通計画におけるパネル調査データの適用性に関する研究, 平成 5・6 年度科学研究費補助金総合研究 (A) 研究成果報告書第 5 章, 飯田恭敬編, pp.31-60.

伊東誠 (1984) 全国幹線旅客純流動の調査方法の研究, 土木計画学研究・論文集, No.1, pp.35-42.

内田敬, 飯田恭敬, 松下晃 (1992) 通勤ドライバーの出発時刻決定行動の実証分析, 土木計画学研究・論文集, No.10, pp.39-46.

佐野紳也 (1990) 質的選択分析—理論と応用, (財) 三菱経済研究所.

杉恵頼寧, 藤原章正, 山根啓典 (1993) 選好意識パネルデータに潜在する消耗バイアスの修正, 土木計画学研究・論文集, No.11, pp.167-174.

鈴木聡, 毛利雄一, 中野敦, 原田昇 (1990) パネルデータに基づく交通手段選択行動の分析, 土木計画学研究・論文集, No.6, pp.537-542.

瀬戸公平, 大塚祐一郎, 飯田克弘, 北村隆一 (1994) 時間利用データを用いた個人の活動に関する分析, 土木学会第 49 回年次学術講演会講演概要集第 4 部, pp.836-837.

新谷洋二, 高橋賢悟, 石橋敏夫, 宮本成夫 (1990) 千葉都市モノレール開通影響調査について, アーバンインフラ・テクノロジー推進会議, 第 2 回技術研究発表論文集, pp.59-65.

西井和夫, 近藤勝直, 古屋秀樹, 鈴木隆 (1995-1) パネルアトリションを考慮した買物場所選択モデル, 土木計画学研究・講演集, No.17, pp.39-42.

西井和夫, 近藤勝直, 古屋秀樹, 鈴木隆 (1995-2) パネルアトリッションを考慮した買物場所選択モデル: 甲府買物パネルデータを用いて, 土木計画学研究・論文集, No.12, pp.389-396.

藤原章正, 杉恵頼寧 (1992) パネルデータを用いた新交通システムに対する選好意識の時間的変化の分析, 都市計画論文集, No.27, pp.397-402.

藤原章正, 杉恵頼寧, 張峻屹 (1996) Mass Point 手法による交通機関選好モデルの消耗及び回答バイアスの修正, 土木計画学研究・論文集, No.13, pp.587-594.

森川高行, 田中小百合 (1993) トリップ分析による郵送方式パーソントリップ調査の適用性に関する研究, 土木計画学研究・講演集, No.16(1), pp.305-310.

山形耕一 (1995) 非集計行動モデルのための調査とデータ作成, 土木学会編, 非集計行動モデルの理論と実際第 4 章, pp.91-122.

協本和昌 (1984) 統計学—見方・考え方, 日本評論社.

Ben-Akiva, M. and Lerman, S. R. (1985) Discrete Choice Analysis, Theory and Application to Travel Demand, MIT Press.

Brownstone, D. and Chu, X. (1992) Multiply Imputed Sampling Weights: A Simple but General Method for Consistent Inference with Panel Attrition Attrition, Paper for the First U.S. Conference on Panels for Transportation Planning.

Brownstone, D. and Golob, T. F. (1992) The Effectiveness of Ridesharing Incentives: Discrete-choice Models of Commuting in Southern California, Regional Science and Urban Economics, Vol. 22, pp.5-24.

Golob, T. F. (1990) The Dynamics of Household Travel Time Expenditures and Car Ownership Decisions, Transportation Research, Vol.24 (A), pp.443-463.

Golob, J. M., Schreurs, L. J. M. and Smit, J. G. (1986) The Design and Policy Applications of a Panel for Studying Changes in Mobility Over Time, Behavioural Research for Transportation Policy, pp.81-95.

Goodwin, P. B. (1986) A Panel Analysis of Changes in Car Ownership and Bus Use, Traffic Engineering and Control, Vol.27, No.10, pp.519-525.

Hausman, J. A. and Wise, D. (1979) A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preference, Econometrica, Vol.46, pp.403-426.

Heckman J. (1976) The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent variables and a Simple Estimator for Such Models, Annals of Econometric and Social Measurement, No.5, pp.475-492.

Hensher D. A. (1987) Issues in the Pre-Analysis of Panel Data, Transportation Research, Vol.21 (A), pp.265-285.

Hensher, D. A., Smith, N. C., Milthorpe, F. M. and Barnard, P. O. (1992) Dimensions of Automobile Demand: A Longitudinal Study of Automobile Ownership and Use, In Study in Regional Science and Urban Economics, Anselin, L. et al (eds.), Vol.22, North Holland, pp.52-86.

Iida, Y., Uchida, T. and Nakahara, M. (1992) Panel Survey on the Effects of Travel Time Information System in Osaka, Paper for the First U.S. Conference on Panels for Transportation Planning.

Kitamura, R. and Bovy, P. (1987) Analysis of Attrition Biases and Trip Reporting Errors for Panel Data, Transportation Research, Vol.21 (A), pp.287-302.

Kumar, A. and Replogle, M. (1992) Low-Cost Trip Purpose Panel Survey: Case Study of Montgomery County, Maryland, Paper for the First U.S. Conference on Panels for Transportation Planning.

Manski, C. F. and Lerman, S. R. (1977) The Estimation of Choice Probabilities from Choice-Based Samples, Econometrica, pp.1977-1988.

Meurs, H. and Ridder, G. (1992) Attrition and Response Effect in the Dutch Mobility Panel, Paper for the First U.S. Conference on Panels for Transportation Planning.

Morikawa, T. (1994) Correlating State Dependence and Serial Correlation in the RP/SP Combined Estimation Method, Transportation, 21, pp.153-165.

Murakami, E. and Ulberg, C. (1992) Current Status of the Puget Sound Transportation Panel, Paper for the First U.S. Conference on Panels for Transportation Planning.

Murakami, E. and Watterson, W. T. (1990) The Puget Sound Transportation Panel after Two Waves, Transportation, Vol.19, pp.141-158.

Pendyala, R. M. and Kitamura, R. (1992) Weighting Methods for Choice Based Panel with Attrition, Paper for the First U.S. Conference on Panels for Transportation Planning.

Uhlman, C. J. and Kim, S. (1992) Designing and Implementing a Panel Study of Commuter Behavior: Lessons for Future Research, Paper for the First U.S. Conference on Panels for Transportation Planning.

van Wissen, L. J. G. and Meurs, H. J. (1989) The Dutch Mobility Panel: Experiences and Evaluation, Transportation, Vol.16, pp.99-119.

Wingley, N., Guy, C., Dunn, R. and O'Brien, L. (1985) The Cardiff Consumer Panel: Methodological Aspects of the Conduct of a Long-term Panel Survey, Transportation Instruments of British Geographers N. S. 10, pp.63-76.

Winner, R. (1983) Attrition Bias in Econometric Models Estimated with Panel Data - The Gary Income Maintenance Experiment, *Econometrica*, Vol.47, pp.177-186.

## 第6章 研究のまとめと今後の課題

非集計分析がわが国に本格的に紹介されて以来、交通行動分析に限らず、非集計モデルを用いた研究は、わが国の土木計画学研究の中で、大きなシェアを占めてきていた。その背景には、社会資本の整備の進展に伴って、政策のより詳細な効果を知る必要性がでてきたことなどの、交通計画におけるニーズの面と、小サンプルでのモデル構築が可能であるという利点と、効用最大化原理という明快な理論的背景による、あらゆる局面での適用のしやすさといったモデルの特性による面が存在する。これらが相乗的に反映した結果として、理論的背景を無視して安易に非集計分析を適用した例も見られ、それらの反動として非集計分析に対する批判も増えた。本研究では、そのような反省のもと、1章でこれまでの非集計分析の研究の流れをレビューし今後の発展の方向を見極めた上で、非集計分析の更なる進展を目指して、1章で示した消費者意思決定過程のパスダイヤグラムに基づいた交通行動のモデリングを行うことを目的とした。具体的には、以下の3つの分析フレームで研究を行った。

- 1) 非集計行動モデルに、主観的・潜在的な要因を取り込んだモデルの構築
- 2) 個人の異質性を考慮した非集計モデルの構築
- 3) パネルデータを用いて、誤差項の系列相関を考慮した非集計行動モデルの構築

これらの分析フレームは、いずれも1960年代後半からの交通行動の意識データ活用法の展開及び1980年代からの動的分析に端緒を發した研究の流れ(Pas (1990))の中にあり、いまだ実用化がされていない交通行動分析の重要課題である。以下、上記の分析フレームに沿ってそれらの研究成果をまとめ、本研究で得られた結論を述べる。

- 1) 非集計行動モデルに、主観的・潜在的な要因を取り込んだモデルの構築

交通行動モデルに主観的・潜在的な要因が重要であるとの指摘は、古くから行われていたが、非集計モデルの背景となるランダム効用理論と整合性を保ち、かつ理論的背景が明確な分析手法が提案されていなかった。そこで本研究では、マーケティングリサーチの分野で用いられる、消費者意思決定構造のパスダイヤグラムをベースとして、近年普及してきた多変量解析手法である線形構造方程式モデル(LISRELモデル)を用いて、快適性、利便性という2つの知覚値を非集計モデルの説明変数として用いたモデルを提案した。このとき、同じ分析フレームで以下の4つの推定方法を試みた。

- ・ LISREL モデルと非集計モデルの段階的推定法

- ・ 認知的不協和を考慮した段階的推定法

- ・ LISREL モデルと離散型選択モデルの同時推定法

- ・ 知覚値指標の離散性を考慮した同時推定法

その結果として、4つの推定法から得られた主観的・潜在の変数は、いずれも選択モデル中で有意なパラメータを持ち、その有効性が確かめられた。ただし、各推定方法の間では線形構造方程式モデルの推定パラメータに違いが見られた。その違いは主に段階推定と同時推定の違いに依存し、複雑な MIMC モデルの推定に対する統計的性質の違いが原因と思われる。段階推定と比較して、統計的性質・データの性質上の観点からより優れていると思われる同時推定モデルが、線形構造方程式モデル部だけを取り上げた場合の適合度が悪くなるなど、それぞれの推定方法には一長一短があり、優劣を判定することはできなかった。ただし、認知的不協和の影響を考慮した段階推定によって得られた結果は、認知的不協和解消行動による知覚値指標のバイアスの存在を示唆しており、このバイアス解消をモデルに組み込む必要性があることが確認された。(2章)

## 2) 個人の異質性を考慮した非集計モデルの構築

交通行動分析に用いられる非集計モデルは、一般的に母集団を同一の効用パラメータを持つ均質な集団と仮定しているが、近年、非集計モデルの分析対象の変化に伴い、個人の異質性を明示的に考慮する必要性が高まってきている。そこで本研究では以下に挙げる5つのアプローチによって、非集計モデルに個人の異質性を取り込むための手法の開発を行った。

- ・ 複数S Pデータを用いて個人ごとにモデルを推定し、得られた個人のパラメータを基準としたセグメンテーションによるアプローチ。

- ・ アンケート調査によって得られた態度指標を用いて、個人の潜在的態度を推計し、それを用いたセグメンテーションを行い、複数S Pデータに適用するアプローチ。

- ・ 潜在的態度変数を用いた潜在セグメントモデルを複数S Pデータに適用するアプローチ。

- ・ 同じ客観的な属性を用いて、アプリアリセグメントと潜在セグメントモデルをパネルデータに適用し、その比較を通じてそれぞれの特長を探るアプローチ。

- ・ 線形構造方程式モデルによって得られた潜在的態度を用いて潜在セグメントモデルをパネルデータに適用するアプローチ。

このうち、複数S Pデータを用いた個人パラメータによるセグメンテーションは、得られた各セグメント内ではパラメータは均質であり、セグメント間では互いに異なることが統計的に検定された。しかし、個人の帰属するセグメントが、実際に利用した交通機関に依存していることが確認され、S Pデータの収集方法に原因を持つバイアスの存在が否定できないという結果になった。よってこのアプローチを採用する場合には、S Pデータ収集法の検討を行う必要があると言える。(3章)

潜在的態度変数に基づくセグメンテーションを、複数S Pデータに適用したアプローチでは、各セグメント内でパラメータは均質でかつセグメント間では互いに異なる、セグメントを規定する潜在的態度変数と整合性のある潜在セグメントが得られた。このアプローチは、先の個人パラメータアプローチと比較して、データ収集が容易で、統計的性質にも問題がないなどの優れた点が存在するが、セグメンテーションに用いる潜在的態度変数と、分析対象となる非集計モデルに用いた説明変数に関連性がない場合には、そのセグメントの評価が困難になるという問題点が存在する。(3章)

先に得られた潜在的態度変数を用いて、潜在クラス分析の応用である潜在セグメントモデルを複数S Pデータに適用するアプローチでは、セグメントを規定する潜在変数から予想されるような解釈の容易な潜在セグメントが得られ、モデル自身の適合度も非常に高いものとなった。ただし、このアプローチは、推定するための尤度関数が複雑となり、得られた推定値が定義域全体での最大の尤度をもたらず点である保証はなく、セグメントを規定する潜在変数を、非集計モデルの説明変数と関連性のないものを用いた場合には、得られるセグメントに整合性が無いなどの問題点はある。しかし、このアプローチは個人ごとのパラメータの違いを構造化したものとなっているため、個人からそれほど多くの情報を得る必要性がなく、個人の異質性を考慮するためには有効な手法と考えられる。(3章)

複数S Pデータを用いてその有効性を確認した潜在セグメントモデルを、同じように同一個人から複数の観測を繰り返すことを前提とするパネルデータに適用した。ここではまず、同一の客観的変数を用いて、アプリアリセグメントと潜在セグメントモデルを推定した。その結果、両



者の全体としての自由度修正済みの適合度は、同じ量の情報を用いたにも関わらず、潜在セグメントモデルの方が高くなった。潜在セグメントモデルによって抽出された潜在セグメントは、アプリオリセグメントと同様に解釈の容易なものとなり、パネルデータに対しても潜在セグメントモデルの有効性は確認できた。（４章）

先に複数S Pデータを用いて行ったものと同様のアプローチで、潜在的態度変数を用いた潜在セグメント分析を、パネルデータを用いた個人の異質性を考慮する事例研究の２つめとして行った。その結果、セグメントを規定する潜在的態度変数と矛盾しない潜在セグメントが抽出され、適合度も非常に高い推定結果が得られた。（４章）

### ３）パネルデータを用いて、非集計行動モデルに時間的要素を取り込んだモデルの構築

パネルデータは個人の時間的な変化による特性を含むことができ、交通行動を動的に分析するためのソースとして、近年交通行動分析の分野で非常に研究が盛んである。本研究では、パネルデータを個人の異質性を考慮したモデルに、同一個人から得る情報量の多さというパネルデータの特長を生かして適用したが、パネルデータには、同一個人から繰り返しデータを採取することに伴う欠点も存在する。それが、系列相関の問題であり、パネル消耗バイアスの問題である。パネルデータの欠点は他にもいくつかあるが、本研究ではこの２点に的を絞り、パネルデータを用いたモデル構築を行った。

パネルデータは同一個人に時点を変えて観測を行うため、非観測要因が時間的に安定した構造を持っていた場合、独立性と同一性を仮定する非集計モデルの誤差項に、同一個人内で時系列的な相関関係が生じてしまう。この相関関係は、個人ごとに異なるため、個人の非観測異質性の問題ともリンクしてくるものである。そこで本研究では、同一個人の誤差項を時点間で共通なものと時点間で独立なものに分解する手法を採用し、系列相関を考慮した非集計分析を行った。パネルデータを用いて、潜在セグメントモデルと同時に系列相関を考慮した事例研究の結果からは、系列相関を考慮する・しないに関わらず、潜在セグメントモデルの推定パラメータにそれほど差が無かった。つまり、時間的に安定した構造を持つ個人の非観測異質性が効用パラメータに与える影響は、それほど大きくないことが確認された。（４章）

パネルデータには、同一個人から一定の間隔を置いて繰り返し観測を行うため、サンプルが徐々に減少して行くというパネル消耗の問題がある。これもサンプルの非観測要因（異質性）が、

選択とパネル滞在の双方に影響することによって引き起こされる問題である。そこで本研究では、その非観測異質性を系列相関項として考慮したモデル構築を行った。このとき、パネル消耗の問題とパネル調査に限らない問題である調査の初期回収時の無回答を、同一の行動原理に基づく行動であると仮定し、調査への不参加行動を包括的に取り扱うフレームで分析した。この仮定を導入することで、母集団の構成が不明な場合でも、母集団での属性分布を推定することができ、各サンプルの母集団と比較したサンプルウェイトを求めることができる。パネルデータを用いて事例研究を行った結果では、本モデルは消耗を規定する変数がそれほど有効でなかったことと、調査の回収率を情報として用いない構造のモデルとなっていたため、期待していたほどの効果は得られなかった。（５章）

本研究では、当初の研究目的に添って３つの分析フレームに基づいて研究を行った。その結果、「非集計行動モデルに主観的・潜在的な要因を取り込んだモデルの構築」「個人の異質性を考慮した非集計モデルの構築」については、本研究で提案したモデルの有効性と現実的な適用性はいくつかの事例研究を通じて確認でき、方法論的に一定の成果を残すことができたと考える。ただし、知覚値・態度データやS Pデータなどのデータそのものの特性について本研究では深く検証を行っていない。本研究の分析フレームが実用的に用いられるためには、２章で示したような、認知的不協和バイアスに限らず、これら様々な意識データの、採取方法や質問設定によるバイアスの性質を明確にする必要がある。そのためには、今回用いたものとは別のデータを用いた事例研究を積み重ねる必要がある。また、「パネルデータを用いた、動的な非集計行動モデルの構築」に関しては、５章で行ったパネル消耗バイアスのモデル論的研究が、提案したモデルに改善の余地を多く残すものとなり、今後パネル消耗モデルの改良に重点を置いた更なる研究が必要である。ただ、これまで、消耗バイアスに関する研究に比べて、無回答バイアスに対するモデリングアプローチはほとんど存在していない。本研究はその意味で、無回答バイアスに対するモデリングアプローチの端緒となる足がかりを残せたと言える。

ここで、本研究での成果をベースとした今後の研究課題を３点指摘して、本研究を閉じる。

- １） 本研究で提案した新たなモデルについての予測能力の評価は、主に非集計分析で用いられるMcFaddenの $\bar{p}^2$ （Ben-Akiva and Lerman（1985））を用いて行った。しかし、これの統計的性質はいまだ明確でない。そこで、近年コンピュータの処理能力の上昇に伴って実用的になってきた、マイクロシミュレーションを用いたモデルの予測能力の評価を行うことが、

今後必要となる研究課題である。

- 2) 本研究は各章の成果が、十分他の章に反映されているとはいえない。そこで、各章での成果をもとに、意思決定構造を明示的に取り扱った動的な交通行動モデルシステムの包括的な構築が望まれる。この研究を始めた頃に比べて、この数年のコンピュータの処理能力の飛躍的上昇に伴い、モデル推定・予測の計算負荷は十分実用的な水準に達しつつあるため、包括的モデルシステムの構築が今後の研究の課題とされる。
- 3) 1章で提示した消費者意思決定構造のパスダイヤグラムは、その一部を提示したもので、他にも意思決定にはいくつかの要因が影響を与えている。例えば、選択肢集合の違いが選択に与える影響である。1章で指摘したように、選択肢集合形成過程に関する研究は、いまだ発展途上である。本研究で提案した個人の異質性を考慮したモデルや潜在的要因を考慮したモデルは、選択肢集合の異質性にも適用が可能なものであり、この研究で得られた知見を選択肢集合の異質性の問題に適用することは十分実用的であり、今後の研究課題として提示できる。

## 第6章 参考文献

- Pas, E. (1990) Is Travel Demand Analysis and Modeling in the Doldrums?, Jones, P. (ed.), Developments in Dynamic and Activity-Based Approaches to Travel Analysis, Avebury, pp.3-27.
- Ben-Akiva, M. and Lerman, S. R. (1985) Discrete Choice Analysis, Theory and Application to Travel Demand, MIT Press.

## 謝 辞

本研究の遂行にあたっては、多くの方々のご指導とご協力を賜った。ここに記して感謝の意を表す次第である。

まず、本研究の遂行ならびに本論文の作成にあたり、終始適切な指導を賜った名古屋大学大学院工学研究科助教授の森川高行先生に対して深甚たる感謝の意を表したい。筆者が京都大学の学部生時代より、交通行動分析の分野で新進気鋭の研究者である同助教授に、研究面に限らず様々なことを学ばせて頂けたことは、筆者の最も大きな幸いの一つであった。

次に、名古屋大学において自由な研究の場を与えて頂いた、名古屋大学大学院工学研究科教授の河上省吾先生に感謝の意を表したい。また、京都大学大学院工学研究科教授の北村隆一先生には、本論文を作成するにあたって非常に丁寧なご指導を賜った。両先生には本論文にまとめられた研究について幾度となく御助言を頂いた。また筆者が学生時代にご指導を賜った現神戸大学工学部教授の黒田勝彦先生、現立命館大学理工学部教授の塚口博司先生、現岡山大学環境理工学部講師の谷口守先生、現大阪大学工学部助手飯田克弘先生には、筆者が非常に怠惰な学生であったにもかかわらず、常に暖かな励ましとご助言を頂いた。そして、山梨大学工学部助教授の西井和夫先生には、本研究で用いたデータを快く提供していただき、かつ研究に対するご助言を頂いた。これら諸先生方に感謝の意を表したい。

また、本研究を進めるにあたって、貴重な発表の場や有益なご助言を頂いた、名古屋交通問題研究会の諸先生方に感謝の意を表したい。

最後に研究を進めるにあたり、京都大学工学部交通土木工学教室起終点施設学講座（現 京都大学大学院工学研究科土木システム工学専攻社会システム工学講座交通システム分野）および名古屋大学工学部土木工学教室土木計画学講座（現 名古屋大学大学院工学研究科社会資本計画学講座）の諸先輩方、諸兄には多くの貴重な意見やご協力を頂いた。特にデータ収集・整理、統計計算、論文作成にあたっては山田菊子氏（(株)三菱総合研究所）、杉本直氏（岐阜県庁）、岡崎真人氏（東海旅客鉄道株式会社）、杉山幸司氏（(株)建設技術研究所）には多大なご協力を頂いた。感謝の意を表します。

本研究はここに記しきれない多くの方々の学恩、ご支援を得て成立していることを銘記し、深く感謝いたします。